

1213.43573X00

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicants: WATANABE

Serial No.:

Filed: March 4, 2004

For: Data-Migration Method

Group:

Examiner:

**LETTER CLAIMING RIGHT OF PRIORITY**

Mail Stop: New Appln.  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

March 4, 2004

Sir:

Under the provisions of 35 USC 119, applicant hereby claims the right of priority based on Japanese Patent Application No. 2004-005635, filed January 13, 2004.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundidge  
Registration No. 29,621

CIB/jla  
(703) 312-6600  
Attachment

日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application: 2004年 1月13日

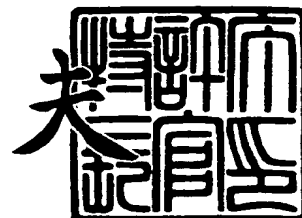
出願番号  
Application Number: 特願2004-005635  
[ST. 10/C]: [JP 2004-005635]

出願人  
Applicant(s): 株式会社日立製作所

2004年 2月 4日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



出証番号 出証特2004-3006183

【書類名】 特許願  
【整理番号】 CM0311024  
【提出日】 平成16年 1月13日  
【あて先】 特許庁長官殿  
【国際特許分類】 G06F 12/00  
【発明者】  
    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所  
                                システム開発研究所内  
    【氏名】 渡辺 直企  
【特許出願人】  
    【識別番号】 000005108  
    【氏名又は名称】 株式会社日立製作所  
【代理人】  
    【識別番号】 100075513  
    【弁理士】  
    【氏名又は名称】 後藤 政喜  
【選任した代理人】  
    【識別番号】 100084537  
    【弁理士】  
    【氏名又は名称】 松田 嘉夫  
【選任した代理人】  
    【識別番号】 100114236  
    【弁理士】  
    【氏名又は名称】 藤井 正弘  
【手数料の表示】  
    【予納台帳番号】 019839  
    【納付金額】 21,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1  
    【包括委任状番号】 0110326

**【書類名】 特許請求の範囲****【請求項 1】**

ホストコンピュータと、前記ホストコンピュータに接続されている記憶サブシステムと、を備えた情報処理システムにおいて用いられ、移行元記憶サブシステムに記憶されているデータを、移行先記憶サブシステムへ移行するデータ移行方法であって、

前記移行元記憶サブシステムから前記移行先記憶サブシステムへデータを移行する前に、経路移行状態が設けられており、

前記経路移行状態では、

前記ホストコンピュータから前記移行元記憶サブシステム及び前記移行先記憶サブシステムへのアクセスを可能とするように経路を設定し、

前記移行先記憶サブシステムは、前記ホストコンピュータからの読み出し要求に対し前記移行元記憶サブシステムからデータを読み出して前記ホストコンピュータに送信し、前記ホストコンピュータからの書き込み要求に対し前記移行元記憶サブシステムにデータを書き込むことを特徴とするデータ移行方法。

**【請求項 2】**

前記経路移行状態において、前記移行先記憶サブシステムは、

前記ホストコンピュータからの書き込み要求に対し前記移行元記憶サブシステムにデータを書き込み、

当該書き込みの確認後、前記ホストコンピュータに対し終了報告を送信することを特徴とする請求項 1 に記載のデータ移行方法。

**【請求項 3】**

前記経路移行状態において、前記移行先記憶サブシステムは、

前記ホストコンピュータからの読み出し要求に対し、前記移行元記憶サブシステムからデータを読み出して、前記ホストコンピュータに送信することを特徴とする請求項 1 に記載のデータ移行方法。

**【請求項 4】**

前記ホストコンピュータから前記移行先記憶サブシステムへのアクセスを禁止し、前記ホストコンピュータから前記移行元記憶サブシステムへのアクセスを可能とするように経路を設定する移行処理前状態が、前記経路移行状態の前に設けられていることを特徴とする請求項 1 に記載のデータ移行方法。

**【請求項 5】**

前記ホストコンピュータから前記移行元記憶サブシステムへのアクセスを禁止し、前記移行先記憶サブシステムへのアクセスを可能とする経路を設定しデータ移行状態が、前記経路移行状態の後に設けられていることを特徴とする請求項 1 に記載のデータ移行方法。

**【請求項 6】**

前記データ移行状態においては、前記移行先記憶サブシステムは、

前記ホストコンピュータからの読み出し要求がデータ移行済み領域に対するものであれば、前記移行先記憶サブシステムからデータ読み出して前記ホストコンピュータへ送信し、

前記ホストコンピュータからの読み出し要求がデータが移行済みでない領域に対するものであれば、前記移行元記憶サブシステムからデータを読み出して前記ホストコンピュータへ送信することを特徴とする請求項 5 に記載のデータ移行方法。

**【請求項 7】**

前記データ移行状態の終了後においても、前記ホストコンピュータから前記移行元記憶サブシステムへのアクセスを禁止し、前記移行先記憶サブシステムへのアクセスを可能とするように経路を設定することを特徴とする請求項 5 に記載のデータ移行方式。

**【請求項 8】**

前記ホストコンピュータ、前記移行元記憶サブシステム及び前記移行先記憶サブシステム間の接続形態を変更すること；

前記ホストコンピュータ、前記移行元記憶サブシステム及び前記移行先記憶サブシステ

ム間を接続するネットワークによるアクセス制限を用いること；

又は、

前記記憶サブシステムによるアクセス制限を用いること；の少なくとも一つによって、前記記憶サブシステムへのアクセスの経路を設定することを特徴とする請求項 1 に記載のデータ移行方法。

【請求項 9】

前記移行元記憶サブシステムから前記移行先記憶サブシステムへデータを移行する前に、前記経路移行状態において設定された経路が正しいか否かを検証する経路検証処理を行うことを特徴とする請求項 1 に記載のデータ移行方法。

【請求項 10】

前記経路検証処理中は、前記ホストコンピュータから前記移行元記憶サブシステムへのアクセスを禁止し、前記移行先記憶サブシステムへのアクセスを可能とするように経路を設定し、

前記移行先記憶サブシステムは、

前記ホストコンピュータからの書き込み要求を前記移行先記憶サブシステムに書き込み、前記移行先記憶サブシステムは該書き込まれたデータを記憶し、

前記ホストコンピュータからの読み出し要求に対し、前記書き込まれたデータを参照して、前記経路検証処理中における前記移行先記憶サブシステムに記憶されているデータの更新の有無を判定し、

前記経路検証処理中に前記移行先記憶サブシステムに記憶されているデータが更新されている場合は、前記移行先記憶サブシステムからデータを読み出して前記ホストコンピュータにデータを送信し、

前記経路検証中に前記移行先記憶サブシステムに記憶されているデータが更新されていない場合は、前記移行元記憶サブシステムからデータを読み出して前記ホストコンピュータにデータを送信することを特徴とする請求項 9 に記載のデータ移行方法。

【請求項 11】

前記経路検証処理において、設定された経路の誤りが発見された場合、前記経路検証中に更新されたデータを破棄し、経路移行前の状態に戻すことを特徴とする請求項 9 に記載のデータ移行方法。

【請求項 12】

前記経路検証処理において、設定された経路の誤りが発見されなかった場合、前記経路検証中に更新されたデータを破棄し、経路移行前の状態に戻してから、データ移行処理を実行することを特徴とする請求項 9 に記載のデータ移行方法。

【請求項 13】

複数のホストコンピュータと、前記複数のホストコンピュータに接続されている記憶サブシステムと、を備えた情報処理システムにおいて用いられ、移行元記憶サブシステムに記憶されているデータを、移行先記憶サブシステムへ移行するデータ移行方法であって、

前記移行元記憶サブシステムから前記移行先記憶サブシステムへデータを移行する前に、経路移行状態が設けられており、

前記経路移行状態では、

前記複数のホストコンピュータは、前記移行先記憶サブシステムへアクセスする移行済ホストコンピュータと、前記移行元記憶サブシステムへアクセスする未移行ホストコンピュータと、が設定され、

前記移行先記憶サブシステムは、

前記移行済ホストコンピュータからの読み出し要求に対応して前記移行元記憶サブシステムからデータを読み出して当該ホストコンピュータに送信し、

前記移行済ホストコンピュータからの書き込み要求に対応して前記移行元記憶サブシステムにデータを書き込み、当該書き込みの確認後、当該ホストコンピュータに対し終了報告を送信し、

前記移行元記憶サブシステムは、

前記未移行ホストコンピュータからの読み出し要求に対応してデータを読み出して当該ホストコンピュータに送信し、

前記未移行ホストコンピュータからの書き込み要求に対応してデータを書き込み、当該書き込みの確認後、当該ホストコンピュータに対し終了報告を送信することを特徴とするデータ移行方法。

**【書類名】 明細書****【発明の名称】 データ移行方法****【技術分野】****【0001】**

本発明は、コンピュータシステムで用いられる記憶サブシステムに格納されたデータを移行する方法に関し、特に、記憶サブシステム群間を接続し、上位装置たる複数のホストコンピュータ群に影響を与えることなく旧記憶サブシステムのデータを新記憶サブシステムに移行する、データ移行方法に関する。

**【背景技術】****【0002】**

記憶システムに格納されたデータの移行には、上位装置（ホスト）からのアクセス中のデータ移行には、拡張リモートコピー機能（XRC：Extended Remote Copy）、又は対等リモートコピー機能（PPRC：Peer-to-Peer Remote Copy）がIBM社から提案されている（例えば、非特許文献1参照。）。

**【0003】**

また、SymmetrixData Migration Service（SDMS）がEMC社から提案されている（例えば、非特許文献2参照。）。

**【非特許文献1】**「Implementing ESS Copy Services on S/390」, IBM P.502.8.5 D ASD migration

**【非特許文献2】**「Symmetrix Data Migration Services」, EMC Corporation, インターネットURL <[http://japan.emc.com/pdf/products/sdms/sdms\\_ds.pdf](http://japan.emc.com/pdf/products/sdms/sdms_ds.pdf)>

**【発明の開示】****【発明が解決しようとする課題】****【0004】**

しかし、これらの従来技術においては、ホストから記憶サブシステムに対するアクセス経路を、旧記憶サブシステムから新記憶サブシステムに切り替える経路切り替え時にはデータ移行にかかわる全てのホストコンピュータを同時に停止し、経路切り替えを行う必要がある。このときデータ移行対象システムが大規模、複雑となると移行時間も増加しシステム全体の可用性が低下する。

**【0005】**

本発明は、大規模、複雑なシステムにおいても可用性を低下させることのないデータ移行方法を提供することを目的とする。

**【課題を解決するための手段】****【0006】**

本発明では、ホストコンピュータと、前記ホストコンピュータに接続されている記憶サブシステムと、を備えた情報処理システムにおいて用いられ、移行元記憶サブシステムに記憶されているデータを、移行先記憶サブシステムへ移行するデータ移行方法であって、前記移行元記憶サブシステムから前記移行先記憶サブシステムへデータを移行する前に、経路移行状態が設けられており、前記経路移行状態では、前記ホストコンピュータから前記移行元記憶サブシステム及び前記移行先記憶サブシステムへのアクセスを可能とするように経路を設定し、前記移行先記憶サブシステムは、前記ホストコンピュータからの書き込み要求に対し前記移行元記憶サブシステムにデータを書き込み、前記ホストコンピュータからの読み出し要求に対し前記移行元記憶サブシステムからデータを読み出して前記ホストコンピュータに送信する。

**【発明の効果】****【0007】**

本発明によると、大規模、複雑なシステムにおいても可用性の低下を抑制することができる。

**【発明を実施するための最良の形態】****【0008】**

以下、本発明の実施の形態を図面を参照して説明する。

【0009】

図1は、本発明の実施の形態の情報処理システムの構成を表すブロック図である。

【0010】

この情報処理システムは、ホストコンピュータ101a、101b、旧記憶サブシステム103a、及び、新記憶サブシステム103bから構成される。ここで、記憶サブシステムとは、上位装置（ホストコンピュータ）に対し情報を授受する制御部（コントローラ201）と、情報を格納するディスク装置210を内蔵する記憶サブシステムである。本実施の形態では、記憶サブシステム103の例としてディスクアレイ装置について説明するが、本発明が適用される記憶サブシステムはディスクアレイ装置に限られず、他の記憶サブシステムを用いるものであってもよい。

【0011】

また、ホスト101、旧記憶サブシステム103a、新記憶サブシステム103bは1以上であれば単数でも複数でもよい。

【0012】

さらに、旧記憶サブシステム103aはデータの移行元であり、新記憶サブシステム103bはデータの移行先である。なお、以下の説明では、新記憶サブシステム103aと旧記憶サブシステム103bとを、符号a、bによって区別するが、新/旧を区別する必要がないときには符号のa、bを省いて記載する。

【0013】

各ホスト101は、ワークステーション、マイクロコンピュータ又はメインフレームコンピュータ等であり、アプリケーションプログラムやデータベースシステムが稼働している。

【0014】

記憶サブシステム103は、一つ以上のインタフェース104を有し、ストレージエリアネットワーク（SAN）102を経由してホスト101と接続されている。本実施の形態は、ホスト101と記憶サブシステム103とがSAN102によって接続されているが、SAN102によらずホスト101と記憶サブシステム103を直接接続してもよい。インタフェース104は、SCSI（Small Computer System Interface）、ファイバチャネル（Fibre channel）、iSCSI（Internet SCSI）やFICON、ESCON等の記憶サブシステム向けインタフェースである。本発明においてはSAN102の構成及びインタフェース104の種類に特に制限はない。以下、本実施の形態の説明においては、インタフェース104の例としてファイバチャネルインタフェースについて説明する。

【0015】

記憶サブシステム103は、一つ以上の論理ボリューム105を有する。ホスト101からはインタフェース104を介して論理ボリューム105への読み出し・書き込み等のアクセスが可能である。このとき各インタフェース104が提供する記憶サブシステム用のプロトコルを使用する。ファイバチャネルではFCP（Fibre Channel Protocol for SCSI）、FICONではFC-SB（Single Byte Protocol）を用いる。本発明において用いられるプロトコルの種別は特に制限はない。

【0016】

新記憶サブシステム103bは、移行処理に必要な移行プログラム106、構成情報107、制御情報108、更新情報109及びデータ110を保持する。

【0017】

移行プログラム106は、移行処理を実行するためのプログラムである。構成情報107及び制御情報108は記憶サブシステム103の構成と制御に関する情報であり、移行プログラム106が該情報を参照し移行処理を実行する。移行プログラム106が使用する制御情報としてはデータ移行処理の進捗を示す進捗ポイントがある。

【0018】

更新情報109は、データ移行処理時にホスト101から受け付けた書き込み要求に関



する更新情報を保持し、後述するビットマップ等の形態をとることができる。記憶サブシステム 103 には、ホスト 101 からのデータの読み出し・書き込み要求に関するデータ 110 を一時的に記憶するキャッシュメモリが設けられており、通常動作時のデータ入出力処理の高速化、データ移行処理のデータ処理を実施する（図 9 の 905、図 11 の 1105 等）。

#### 【0019】

移行プログラム 106 は、論理ボリューム 105 毎に設けられた進捗ポインタとビットマップを用い、ボリューム 105 毎にデータ移行処理を行う。

#### 【0020】

なお、ここまで説明した各プログラム 106 は、コンパクトディスクや光磁気ディスクといった可搬記憶媒体を用いて、あるいは、管理ネットワーク 111 を介して、他の装置から各記憶サブシステム 103 に設けられる記憶媒体にインストールされる。

#### 【0021】

移行管理ホスト 112 は、ワークステーション、パーソナルコンピュータ又はメインフレーム等のコンピュータであり、CPU、メモリ及び記憶装置を備えている。CPU では移行管理プログラム 113 が動作しており、旧記憶サブシステム 103a から、新記憶サブシステム 103b への、データ移行処理を管理する。また、メモリ（又は、記憶装置）には、移行管理テーブル 114 が保持されている。

#### 【0022】

移行管理プログラム 113 は、データ移行処理を監視・管理するものである。移行管理プログラム 113 は、該移行管理テーブル 114 を使用してデータ移行処理の監視、制御を行う。具体的には、移行管理プログラム 113 は、管理用ネットワーク 111 を介して接続された各記憶サブシステム 103 の状態の監視と、構成変更、データ移行処理等を実施する。一般に管理ネットワーク 111 は IP プロトコルを用いたイーサネット（登録商標、以下同じ）等にて構成されているが、どのようなネットワークでもよい。

#### 【0023】

また、移行管理プログラム 113 は、移行管理テーブル 114 の全体又は一部を新記憶サブシステム 103b の移行プログラム 106 に送付する。移行プログラム 106 は、送付された移行管理テーブル 114 に基づいて、構成情報 107 及び制御情報 108 を作成する。

#### 【0024】

本実施の形態では、ホスト 101、記憶サブシステム 102 とは別に移行管理ホスト 112 を設けているが、移行管理プログラム 113、移行管理テーブル 114 は移行処理を行うシステム内のどの装置に実装することができる。

#### 【0025】

図 2 に、本発明の実施の形態の記憶サブシステム 103 の構成図である。

#### 【0026】

各記憶サブシステム 104 は、コントローラ 201、一つ以上のディスク装置 210 を有する。コントローラ 201 は、ホストアダプタ 203、メモリ 205、ディスクアダプタ 207、プロセッサ（CPU）202 及びネットワークコントローラ 204 を有する。各々の構成の数は、本発明の本質とは関係ないが性能及び信頼性の観点から多重化することが好ましい。

#### 【0027】

ホストアダプタ 203 は、ファイバチャネル等のインタフェース 104 に関するプロトコルを制御する。

#### 【0028】

ネットワークコントローラ 204 は、監視系ネットワークのプロトコル制御を行い、移行管理ホスト 112 との通信を実行する。

#### 【0029】

メモリ 205 には、データ移行処理に関するプログラム及びデータが格納されている。

具体的には、データ移行処理を実現するための移行プログラム106、データ移行処理の構成情報107、制御情報108、更新情報109及びデータ110が格納される。メモリ205には、これらのプログラム及びデータの他に、記憶サブシステム103を制御するために必要な制御プログラムや制御情報、ホスト101に対する入出力処理に関するキャッシュデータ110が格納される。高信頼化のために、メモリ205を二重化したり、メモリ205への電源を二重化することが好ましい。

#### 【0030】

ディスクアダプタ207は、ホストアダプタ203と同様に、ファイバチャネル等のディスクインタフェース209に関するプロトコル処理を行う。

#### 【0031】

ディスク装置210は、ディスクインタフェース209を介してコントローラ201からの読み出し、書き込み等のコマンドを受け付け、コマンドによって規定される処理を行う。高信頼化のため、ディスクインタフェース209を二重化することが好ましい。

#### 【0032】

記憶サブシステム103は、ディスク装置210を複数組み合わせ冗長構成とし、この中に論理的なデバイス（論理ボリューム）105を作成する。プロセッサ（CPU）202は、記憶サブシステム103に関する処理を実行する。プロセッサ202は内部バス208を介しコントローラ201内のホストアダプタ203、ディスクアダプタ207及びネットワークコントローラ204と接続され、プロセッサ202がこれらを制御する。また、プロセッサ202は、メモリ205とも内部バス208を介し接続されており、メモリ205に格納されたプログラム106や制御情報108をロードして実行する。旧記憶サブシステム103aは新記憶サブシステム103bと同様に図2に示す構成を有する。なお、旧記憶サブシステム103aはデータ移行処理に対応する必要はなく、ホスト101に対するデータ入出力を提供する機能のみを有すれば足りる。

#### 【0033】

また、本実施の形態においては簡素な内部構成を用い説明したが、同等機能を実現可能な場合は内部構成に関しては特に制約はない。例えば、特開平10-333836号公報に示すように、内部バス208の代わりにスイッチを使用してコントローラの各構成間で通信するように構成してもよい。

#### 【0034】

図3は、本発明の実施の形態の移行処理のフローチャートであり、図4は、図3に示す移行処理における情報処理システムの動作の説明図である。

#### 【0035】

図4において、データ移行処理は、移行先記憶サブシステム103bの移行プログラム106によって行われ、移行元記憶サブシステム（旧記憶サブシステム）103aに格納されたデータが移行先記憶サブシステム（新記憶サブシステム）103bへ移行される。

#### 【0036】

まず移行前の状態（図4（a））において、移行管理プログラム113はデータ移行のための初期設定をする。具体的には、移行元記憶サブシステム103aの論理ボリューム105aと移行先記憶サブシステム103bの論理ボリューム105bのペアを設定する。移行のためのペアは論理ボリューム105単位で設定する。この際関連するホスト101（ホスト101で稼動するアプリケーション、ファイルシステム等の情報）も設定する（301）。

#### 【0037】

初期設定の後、移行先記憶サブシステム103bをSAN102に接続し、情報処理システムに追加する（302）。このとき、移行先記憶サブシステム103bは物理的にSAN102に接続されているものの、ホスト101からのアクセスを受け付けないように、移行管理プログラム113が移行先記憶サブシステム103bに対して指示をする（図4（b））。

#### 【0038】

次に経路移行処理(303)を開始する。経路移行処理では移行先記憶サブシステム103bを経路移行状態とする。移行管理プログラム113は移行管理テーブル113によって管理される全ての移行先記憶サブシステム103bに対し、以下の動作を行うための指示をする。

【0039】

経路移行中は、経路移行を終えたホスト101aは移行先記憶サブシステム103bへアクセスする。このとき、ホスト101aからの読み出し要求を受信すると、移行先記憶サブシステム103bは当該リード要求に対応するデータを移行元記憶サブシステム103aから読み出し、ホスト101aへ返送する。また、ホスト101aからの書き込み要求を受信すると、更新すべきデータを移行元記憶サブシステム103aへ書き込む(図4(c))。

【0040】

一方、経路移行を行っていないホスト101bは移行元記憶サブシステムへアクセスする(図4(c))。ホスト101の経路移行はアクセス先記憶サブシステム103を移行元記憶サブシステム103aから移行先記憶サブシステム103bへ切り替える処理である。例えば、ファイバチャネルプロトコルであれば、ホスト101に設定されている記憶サブシステムの情報(WWN:World Wide Name)を移行元記憶サブシステム103aから移行先記憶サブシステム103bへ変更する。iSCSIであればiSCSI Nameを移行元記憶サブシステム103aから移行先記憶サブシステム103bへ変更する。上記経路移行処理は管理者が手動で構成ファイルを変更してもよいが、ホスト101上の経路管理プログラム(例えば、HDL M:Hitachi Dynamic Link Manager等)と移行管理プログラム112が連携することによって実現することもできる。

【0041】

これらの経路移行処理が全て終了すると、全てのホスト101は移行先記憶サブシステム103bにアクセスし、移行元記憶サブシステム103aは、ホストからアクセスされなくなる(図4(d))。

【0042】

経路移行処理が終了するとデータ移行処理を実行する(304)。移行管理プログラム113は移行先記憶サブシステム103bが移行元記憶サブシステム103aのデータを転送するよう移行先記憶サブシステム103bに指示する。

【0043】

全てのデータの転送が終了すると、移行元記憶サブシステム103aをSAN102から切り離し、情報処理システムから削除する(図4(f))。これで、移行元記憶サブシステム103aにフォーマット処理等を施して、格納されていたデータを消去した後に、他のシステムに組み込み、別の目的に使用することができる。

【0044】

上記の処理では、移行元記憶サブシステム103a、移行先記憶サブシステム103bは、それぞれの状態において動作が異なる。各状態の管理と各記憶サブシステム103に対する動作の指示は移行管理プログラム113が行う。

【0045】

図5に、本発明の実施の形態の移行処理の各状態における移行元記憶サブシステム103a、移行先記憶サブシステム103bの動作を示す。

【0046】

本実施の形態においては、移行処理状態(移行ステータス)は「移行処理前」(505)、「経路移行状態」(506)、「データ移行状態」(507)、「移行処理後」(508)の4種類がある。前述した移行処理フローチャート(図3)において、処理301、302の実行中が「移行処理前」505で、処理303の実行中が「経路移行状態」506で、処理304の実行中が「データ移行状態」507で、処理304の実行終了後が「移行処理後」508となる。

【0047】

図5において、カラム501には各状態に対する番号を示し、カラム502には移行元記憶サブシステム103aの各読み出し・書き込み処理に対する動作を示し、カラム504には移行先記憶サブシステム103bの各読み出し・書き込み処理に対する動作を示す。

#### 【0048】

移行元記憶サブシステム103aは、データ移行処理実行前（移行処理前（505）、経路移行状態（506））は、記憶サブシステム103の通常状態における動作と同じ動作をする。すなわち、ホスト101は、移行元記憶サブシステム103aに対する読み出し・書き込み処理を行う。また、移行元記憶サブシステム103aは、データ移行処理実行後（データ移行状態（507）、移行処理後（508））には、ホスト101から移行元記憶サブシステム103aに対する読み出し・書き込み処理等によるアクセスができない状態にする。

#### 【0049】

移行先記憶サブシステム103bは、移行処理前（505）には、ホスト101からデータの読み出し、書き込み処理等によるアクセスができない状態にする。

#### 【0050】

移行先記憶サブシステム103bは、経路移行状態（506）に書き込み要求を受信すると、移行先記憶サブシステム103bで書き込みデータを一旦受け取った後に、移行元記憶サブシステム103aに該データを書き込む。また、読み出し要求を受信すると、最新のデータが移行元記憶サブシステム103aに格納されているため、移行先記憶サブシステム103bは移行元記憶サブシステム103aの該当するアドレスからデータを読み出し、ホスト101へ要求データを送信する。すなわち、ホスト101から移行先記憶サブシステム103bに対する読み出し・書き込み要求に応じて移行元記憶サブシステム103aが最新の状態となるように、他方の記憶サブシステム103が動作する。

#### 【0051】

データ移行状態（507）では、移行先記憶サブシステム103bは、移行元記憶サブシステム103aから移行先記憶サブシステム103bへのデータ移行処理を行う。よって、データ移行が済んでいない領域に対する読み出し要求は、移行元記憶サブシステム103aからデータを読み出す。しかし、データ移行が済んでいる領域に対する読み出し要求や、書き込み要求は、移行先記憶サブシステム103aに対して処理される。

#### 【0052】

移行処理後（508）では、移行先記憶サブシステム103bに最新のデータが格納されるため、ホスト101から移行先記憶サブシステム103bに対する読み出し、書き込み要求は移行先記憶サブシステム103bに対して行われる。

#### 【0053】

図6は、本発明の実施の形態の移行管理テーブル114の説明図である。

#### 【0054】

移行管理テーブル114は、移行管理プログラム113によって作成され、移行管理ホスト112に格納される。移行管理プログラム113は移行管理テーブル114によって移行処理を管理し遂行する。移行処理においては、「移行グループ」と呼ばれる管理単位が設けられている。移行グループとしては論理ボリューム105を最小単位としホスト101やファイルシステム、アプリケーション、ユーザ、部門、フロア、建物等の最適な単位を、管理単位として用いる。

#### 【0055】

各移行グループには特有の識別子であるグループID601が付される。移行プログラム113は、管理テーブル（図6）の移行状態欄602によって、各移行グループ毎に移行の状態を管理する。各移行グループは移行を行うボリューム毎に、移行元記憶サブシステムを移行元記憶サブシステムID603によって管理し、移行先記憶サブシステムを移行先記憶サブシステムID604によって管理する。移行元記憶サブシステムID603及び移行先記憶サブシステムID604に基づいて、移行グループが属する記憶サブシ

テム 103 と記憶サブシステム 103 内の論理ボリューム 105 を特定することができる。

【0056】

さらに、移行管理テーブル 114 には、各移行グループに関する詳細情報 605 として、例えばディレクトリの情報が格納されている。この詳細情報 605 によって移行処理を管理する管理者が移行グループの内容を容易に認識することができ、移行処理全体のスケジュール、手続き等を決定することができる。

【0057】

移行管理プログラム 113 は、移行管理テーブル 114 に基づいてボリューム単位で移行処理の進行を指定する。また、移行管理プログラム 113 は、図 5 で示した動作状態を移行元記憶サブシステム 103 a 及び移行先記憶サブシステム 103 b に対して指定する。

【0058】

[経路移行処理]

図 7 は、本発明の実施の形態の経路移行処理（図 3 のステップ 303）の詳細な動作を示すフローチャートである。

【0059】

経路移行処理は移行グループ毎に実行される。図 7 は任意の移行グループにおける移行処理の詳細を示す。各移行グループには一つ以上の移行される論理ボリューム 106 のペアが存在する。

【0060】

まず、移行処理プログラム 113 は、移行グループに属しているホスト 101 のうち、経路を移行していないホスト 101 が存在するかを確認する（701）。

【0061】

全てのホスト 101 の経路移行が終了している場合には、経路移行処理を終了し、データ移行処理（図 3 のステップ 304）を実行する。移行グループ内の移行元記憶サブシステム 103 a のボリューム 105 a に対するホスト 101 からの読み出し、書き込み等のアクセスを制限する。アクセス制限には複数の方法があり、その詳細は後述する。この際データ移行のため移行元記憶サブシステム 103 a と移行先記憶サブシステム 103 b の間の通信経路は確保した状態で保持しておく（702）。その後、移行管理テーブル 114 の該当する移行グループについて移行状態を「経路移行状態」から「データ移行状態」に変更する。

【0062】

経路移行が行われていないホスト 101 が存在する場合には、該ホスト 101 について経路移行の処理（706～708）を実行する。

【0063】

まず、経路移行が行われていないホスト 101 の中から、次に経路移行処理を行うホスト 101 を選択する（703）。

【0064】

その後、移行処理が完了していない記憶サブシステム 103 があるか否かを判定する（704）。そして、該ホスト 101 に関し、経路移行処理をすべき移行グループ内の記憶サブシステム 103 を選択する（705）。この際、該ホスト 101 に関連する記憶サブシステム 103 内のボリューム 105 全てについて以下の経路移行の処理（706～708）を実施する。ホスト 101 は経路の変更を行うために、経路切り替えに関する記憶サブシステム 103 との接続を切断する（706）。通常、この記憶サブシステム 103 との接続の切断は、ホスト 101 のオペレーティングシステムで提供されているアンマウント処理によって実現可能であり、前述した HDLM 等の経路管理プログラムを使用しても実現可能である。

【0065】

移行先記憶サブシステム 103 b は移行元記憶サブシステム 103 a と別の記憶サブシ

ステム 103 の識別子を有するので、ホスト 101 は記憶サブシステム 103 に関する設定情報を変更し、経路を変更する (707)。記憶サブシステム 103 の識別子としてはファイバチャネル SCSI (FCP) では WWN、iSCSI では iSCSI Name、FICON ではファイバチャネルの WWN 又はポート ID を使用する。

#### 【0066】

記憶サブシステム 103 は一つ以上の識別子を有し、ホスト 101 は一つの記憶サブシステム 103 に対して複数の識別子を有する場合がある。この場合、移行管理テーブル 114 において指定された移行処理に関連する識別子に関してのみ経路移行処理を行う。

#### 【0067】

ホスト 101 は経路の変更を終えた後に、経路切り替えに関する記憶サブシステム 103 と接続する (708)。通常はホスト 101 のオペレーティングシステムで提供されているマウント処理によって実現可能である。

#### 【0068】

以後、全ての記憶サブシステムについて経路移行処理が終了するまで、該当する全ての記憶サブシステム 103 に対して 704 ~ 708 の処理を繰り返し実行する。全ての記憶サブシステム 103 に関する経路切り替え処理が終了すると、再度、経路移行が終了していないホスト 101 の有無を検索し、移行グループ内の全てのホスト 101 に関する経路移行処理が終わるまで 703 ~ 708 の処理を繰り返し実行する。

#### 【0069】

以上説明した経路移行処理では、記憶サブシステム 103 を一つずつ変更しているが、複数の記憶サブシステムの経路変更処理 (704 ~ 708) を同時に実行してもよい。

#### 【0070】

また、移行管理テーブルに基づいてホスト 101 を選択した後に、該ホストが関連する記憶サブシステム 103 に関する経路移行処理を実行したが、先に記憶サブシステム 103 を選択し、該記憶サブシステム 103 に関連するホスト 101 に関して経路移行処理を実行してもよい。

#### 【0071】

##### [データ移行処理]

図 8 は、本発明の実施の形態のデータ移行処理 (図 3 のステップ 304) の詳細な動作を示すフローチャートである。

#### 【0072】

データ移行処理は、各移行先記憶サブシステム 103 b の移行プログラム 106 によって実行され、各ボリューム 105 に対する処理を示している。

#### 【0073】

まず、データ移行処理のための各種変数を初期化して、初期設定を行う (801)。代表的な変数としては、移行進捗を示す進捗ポインタや移行状態を示す移行ステータス等の制御情報 108、及び更新状態を示すビットマップ等の更新情報 109 を初期化する。なお、本実施の形態のデータ移行処理では、移行領域の管理にビットマップを用いる例を説明する。

#### 【0074】

移行プログラム 106 は、任意の管理単位に 1 ビットのフラグを割り当て、移行の可否を示すビットマップを構成する。管理単位は SCSI 等でよく使用される 1 ブロック (512 Byte) や、任意の大きさ (例えば、1 MByte 等) でもよい。ビットが ON になっている (値が 1 である) 場合には、最新状態になっているため移行が不要とし、ビットが OFF になっている (値が 0 である) 場合には、最新状態ではないので移行が必要とする。移行プログラム 106 はデータ移行処理時の初期化時に全ビットを OFF に設定し、データ移行処理を開始する。

#### 【0075】

移行ステータスはボリュームに関する状態を示し、移行処理の 4 状態 (図 5 に示す「移行処理前」(505)、「経路移行状態」(506)、「データ移行状態」(507)、「

「移行処理後」(508))のいずれかとなる。進捗ポインタはデータ移行の進捗を示し、初期化時に移行プログラム106によってボリューム105の先頭を示すよう初期化される。データ移行処理を実行する毎に移行プログラム106は進捗ポインタを更新する。通常、進捗ポインタはボリューム105のアドレスを示す。例えば、進捗ポインタにデータ移行が進行したバイト数を加算して、アドレスをインクリメントする。

【0076】

移行先記憶サブシステム103bの移行プログラム106は、データ移行中のボリューム105の進捗ポインタを調査し、データ移行処理の状態を確認する(802)。進捗ポインタがボリューム105の末尾を示しているときには該ボリューム105のデータ移行が終了しており、以後、移行プログラム106は該ボリューム105に対するデータ移行処理を行わない。

【0077】

移行プログラム106は、進捗ポインタが示す領域のデータを移行するために、該領域の状態を確認する(803)。該領域が読み出し、書き込み処理等によって使用されていない場合、該領域を確保し次の処理に進む(805)。一方、該領域が読み出し、書き込み処理等によって使用中の場合、移行プログラム106は該領域が使用可能になるまで待機し(804)、該領域が確保(805)できた後に次の処理に進む。

【0078】

移行プログラム106は、移行元記憶サブシステム103aより該領域に記憶されたデータを読み出し(806)、キャッシュメモリ110に格納し、移行先記憶サブシステム103bのディスク装置210に書き込む(807)。

【0079】

移行プログラム106は、移行元記憶サブシステム103aから移行先記憶サブシステム103bへ移行されたデータに関する制御情報を更新する。ビットマップに関しては移行済み領域に対応するビットをONにし(808)、移行済み領域の大きさだけ進捗ポインタを進める(809)。

【0080】

通常キャッシュメモリ110は二重化、不揮発化されているため、必ずしも上記タイミングで移行先記憶サブシステム103bのディスク装置210へ書き込む必要はない。移行データをキャッシュメモリ110に格納した直後にビットマップや進捗ポインタを更新し、次の移行データの処理を実行してもよい。

【0081】

[書き込み処理]

図9は、本発明の実施の形態の経路移行状態とデータ移行状態における書き込み処理の概要を示し、図10は、この書き込み処理の詳細な動作を示すフローチャートである。

【0082】

経路移行状態(図9(a))では、ホスト101から移行先記憶サブシステム103bへ書き込まれるデータ901は、移行先記憶サブシステム103bではなく、移行元記憶サブシステム103aに格納される。すなわち、経路移行状態では、全ての書き込みは移行先記憶サブシステム103bには残さず移行元記憶サブシステム103aに反映される。

【0083】

具体的には、まず、移行先記憶サブシステム103bは、ホスト101から受け取ったデータ901を移行先記憶サブシステム103bのキャッシュメモリ110へ格納する。そして、移行先記憶サブシステム103bは、キャッシュメモリ110に格納したデータを移行元記憶サブシステム103aのディスク装置210の記憶領域に書き込む。そして、移行元記憶サブシステム103aに書き込んだ後に、ホスト101に対し当該書き込み処理の結果904を返送する。経路移行状態においてはデータ移行処理を伴わないので、ビットマップ109は更新されない(903)。

【0084】

データ移行状態(図9(b))では、ホスト101から移行先記憶サブシステム103bへ書き込まれるデータ901は、移行先記憶サブシステム103bに格納される。

【0085】

具体的には、まず、移行先記憶サブシステム103bは、ホスト101から受け取ったデータ901を移行先記憶サブシステム103bのキャッシュメモリ110へ格納する。そして、データが未移行の場合、書き込みの対象となる領域のデータをこの読み出しタイミングで移行する。

【0086】

データ移行処理における管理単位は、ホスト101からのアクセス単位と異なる(通常は、データ移行処理における管理単位はホスト101のアクセス単位よりも大きい)。データ901の書き込み対象となる領域がデータ移行処理の管理単位よりも小さい場合には、移行先記憶サブシステム103bは移行元記憶サブシステム103aから該領域に関する管理単位の分のデータを読み込んで(905)、キャッシュ110に書き込む。

【0087】

そして、キャッシュ110において、移行元記憶サブシステム103aから読み出されたデータと、ホスト101から受け取ったデータ901を合わせて、移行先記憶サブシステム103bのディスク装置210へ書き込む。このディスク装置210への書き込みと共に、データ移行を管理するビットマップ109の該当するビットをONにする(906)。

【0088】

次に、図10を用いて書き込み処理の詳細な動作を説明する。

【0089】

移行先記憶サブシステム103bの移行プログラム106は、ホスト101から要求を受け付けると移行処理の状態を確認する(1001)。移行処理前は、移行先記憶サブシステム103bに対してアクセスすることができないので(図5参照)、移行処理前にアクセス要求を受け付けた場合には、異常であると判定し、ホストにエラーを送信する(1002)。状態が「移行処理前」でなければ、移行先記憶サブシステム103bにアクセス可能なので、正常であると判定し、処理を継続する。

【0090】

移行先記憶サブシステム103bの移行プログラム106は、書き込み要求の内容を解析する(1003)。そして、該書き込み要求によって指示されたアドレスに対応する領域を選択し、キャッシュメモリ110にデータ格納用領域を確保する(1004)。この際、要求に係るデータが使用可能なキャッシュメモリ110よりも大きい場合は、全データが受け取り済みになるまで、1004～1014の処理を繰り返す(1015)。

【0091】

移行プログラム106は、ホスト101から書き込みデータを受け取り、該受け取ったデータをキャッシュメモリ110に確保された領域に格納する(1005)。

【0092】

そして、状態が「経路移行中」であるか否かを判定し(1006)、「経路移行中」である場合は、そのまま移行元記憶サブシステム103aへデータを書き込む(1007)。そして、全てのデータの書き込みが終了すると(1015)、ホスト101へ当該処理の結果を返送し(1016)、書き込み処理が終了する。

【0093】

一方、「経路移行中」でない場合は、「データ移行中」であるか否かを判定し(1008)、「データ移行中」である場合は、進捗ポイント及びビットマップを参照し、書き込みに係る領域が移行済みか否かを判定する(1009)。

【0094】

移行済み領域である場合は、移行先記憶サブシステム103bへデータを書き込む(1014)。すなわち、通常の記憶サブシステム103のデータ書き込み動作と同様の処理を実行する。なお、データ移行処理が終了した場合、全ての領域について移行済みである



ため、移行先記憶サブシステム 103b へデータが書き込まれる。

【0095】

一方、移行済み領域ではない場合は、移行元記憶サブシステム 103a からデータを読み出し（1010）、該移行元記憶サブシステム 103a から読み出されたデータと、ホスト 101 から受け取ったデータ 901 とを、キャッシュメモリ 110 において合わせ、新領域データを生成する（1011）。具体的には移行元記憶サブシステム 103a から読み出したデータに対し、書き込みデータ 901 を該当する箇所に上書きする。

【0096】

そして、移行プログラム 106 は、生成された新領域データを移行先記憶サブシステム 103b のディスク装置 210 に書き込み（1012）、該領域に対する移行済みビットを ON にして、該領域が移行済みであることを示す（1013）。

【0097】

書き込み要求に関する全てのデータに対して、上記処理が実行された後に、移行プログラム 106 はホスト 101 に対して当該処理の結果を送信する（1016）。

【0098】

本実施の形態の書き込み処理において行われるデータ書き込み（1007、1012、1014）はデータの一貫性を保証するためホストへの処理結果返送処理 1016 より先に実行されなければならない。しかし、処理 1012 及び処理 1014 においては移行先記憶サブシステム 103b にデータが書き込まれ、移行先記憶サブシステム 103b のキャッシュ 110 においてデータの一貫性を保証するため、ホストへのステータス返送処理 1016 と同期しないタイミングで実施してもよい。

【0099】

本実施の形態の書き込み処理では、移行元記憶サブシステム 103a から読み出し、書き込みデータ 901 とキャッシュ 110 で合わせて、ディスク装置 210 に書き込むデータ移行方法について説明したが、他にもビットマップを使用することなく、進捗ポイントのみによって管理することもできる。この場合、全ての書き込みデータ 901 は移行先記憶サブシステム 103b のディスク装置 210 に書き込まれるが、進捗ポイントより大きいアドレスの未移行領域への書き込みデータ 901 は、移行元記憶サブシステム 103a にも書き込む。その後、図 8 で示したデータ移行処理によって移行元サブシステム 103a から移行先記憶サブシステム 103b へデータが移行される。

【0100】

〔読み出し処理〕

図 11 は、本発明の実施の形態の経路移行状態とデータ移行状態における読み出し処理の概要を示し、図 12 は、この読み出し処理の詳細な動作を示すフローチャートである。

【0101】

経路移行状態（図 11（a））では、ホスト 101 から移行先記憶サブシステム 103b には最新データが格納されていないので、必ず移行元記憶サブシステム 103a からデータ 1201 を読み出す（1102）。

【0102】

具体的には、移行先記憶サブシステム 103b は、ホスト 101 から要求されたデータ 1101 を移行元記憶サブシステム 103a のディスク装置 210 の記憶領域から読み出して、移行先記憶サブシステム 103b のキャッシュメモリ 110 へ格納する。移行先記憶サブシステム 103b は、ホスト 101 から要求されたデータ 1101 をホスト 101 に送信する。そして、データ 1101 の送信が終了するとステータス 1104 をホスト 101 へ送信する。経路移行状態においては、データの読み出し処理ではデータ移行処理が発生しないので、ビットマップ 109 は更新されない。

【0103】

データ移行状態（図 11（b））では、データ移行状態で、かつ、ホスト 101 から要求されたデータが未移行の場合、読み出しが行われた領域のデータをこの読み出しタイミングで移行する。

**【0104】**

具体的には、データ移行処理における管理単位は、ホスト101からのアクセス単位と異なる（通常は、データ移行処理における管理単位はホスト101のアクセス単位よりも大きい）。データ1101の読み出し対象となる領域がデータ移行処理の管理単位よりも小さい場合には、移行先記憶サブシステム103bは移行元記憶サブシステム103aから該領域に関する管理単位の分のデータを読み出し（1105）、キャッシュ110に書き込む。

**【0105】**

そして、キャッシュ110に格納された要求データ1101をホスト101へ送信する。その後、移行先記憶サブシステム103bは、移行元記憶サブシステム103aから読み出されキャッシュ110に格納されたデータをディスク装置210へ移行データとして書き込む。このディスク装置210への書き込みと共に、データ移行を管理するビットマップ109の該当するビットをONにする（1103）。

**【0106】**

次に、図12を用いて読み出し処理の詳細な動作を説明する。

**【0107】**

移行先記憶サブシステム103bの移行プログラム106は、ホスト101から要求を受け付けると移行処理の状態を確認する（1201）。移行処理前は、移行先記憶サブシステム103bに対してアクセスすることができないので（図5参照）、移行処理前にアクセス要求を受け付けた場合には、異常であると判定し、ホストにエラーを送信する（1202）。状態が「移行処理前」でなければ、移行先記憶サブシステム103bにアクセス可能なので、正常であると判定し、処理を継続する。

**【0108】**

移行先記憶サブシステム103bの移行プログラム106は、読み出し要求の内容を解析する（1203）。そして、該書き込み要求によって指示されたアドレスに対応する領域を選択し、キャッシュメモリ110にデータ格納用領域を確保する（1204）。この際、要求に係るデータが使用可能なキャッシュメモリ110よりも大きい場合は、全データが送信済みになるまで、1204～1013の処理を繰り返す（1214）。

**【0109】**

移行プログラム106は、状態が「経路移行中」であるか否かを判定し（1205）、「経路移行中」である場合は、そのまま移行元記憶サブシステム103aからデータを読み出し（1206）、全てのデータの読み出しが終了すると（1215）、ホスト101へステータスを返送し（1215）、読み出し処理が終了する。

**【0110】**

一方、「経路移行中」でない場合は、「データ移行中」であるか否かを判定し（1207）、「データ移行中」である場合は、進捗ポインタ及びビットマップを参照し、読み出しに係る領域が移行済みか否かを判定する（1208）。

**【0111】**

移行済み領域である場合は、移行先記憶サブシステム103bからデータを読み出す（1014）。すなわち、通常の記憶サブシステム103のデータ読み出し動作と同様の処理を実行する。なお、データ移行処理が終了した場合、全ての領域について移行済みであるため、移行先記憶サブシステム103bからデータが読み出される。

**【0112】**

一方、移行済み領域ではない場合は、移行元記憶サブシステム103aからデータを読み出し（1209）、キャッシュメモリ110に格納し新領域データとする。そして、新領域データを、移行先記憶サブシステム103bのディスク装置210に書き込み（1210）、該領域に対する移行済みビットをONとし、該領域が移行済みであることを示す（1211）。

**【0113】**

移行先記憶サブシステム103bは、新領域データのうちホスト101から要求された

データをホスト 101 に送信する (1213)。

【0114】

読み出し要求に関する全てのデータに対して、上記処理を実行した後に、移行プログラム 106 はホスト 101 に対してステータスを送信する (1215)。

【0115】

本実施の形態の読み出し処理において行われるデータ書き込み (1210) は、データの一貫性を保証するためホストへのステータス返送処理 1215 より先に実行されなければならないが、移行先記憶サブシステム 103b のキャッシュ 110 においてデータの一貫性を保証するため、ホストへのステータス返送処理 1215 と同期しないタイミングで実施してもよい。

【0116】

本実施の形態の読み出し処理では、未移行領域へのホスト 101 からの読み出し要求のタイミングで該領域を移行し、ディスク装置 210 に書き込むデータ移行方法について説明したが、他にもビットマップを使用することなく、進捗ポイントのみによって管理することもできる。この場合、未移行領域への読み出し要求があると、移行先記憶サブシステム 103b はデータ 1101 のみを移行元記憶サブシステム 103a から読み出し、ホスト 101 へ送信し、移行処理は行わない。その後、図 8 で示したデータ移行処理によって移行元記憶サブシステム 103a から移行先記憶サブシステム 103b へデータが移行される。

【0117】

[アクセス制限]

図 13 ～ 図 17 は、記憶サブシステム 103 に対するアクセス制限の例を示している。物理的な接続によってアクセスの可否を制御する場合と、論理的に (アクセス管理によって) 制御する場合とが考えられる。さらに、論理的に制御する場合も、ネットワーク上でアクセス制限を行う場合と、記憶サブシステム側でアクセス制限を行う場合とが考えられる。以下これらの 3 種類の例について実施の形態を示す。

【0118】

[結線による例]

図 13 は、本発明の実施の形態の情報処理システムの、物理的な接続によってアクセス制御を行う動作の説明図である。

【0119】

図 13 に示す実施の形態では、移行元記憶サブシステムの SAN 102a と、移行先記憶サブシステムの SAN 102b と存在する。移行処理は図 5 に示すとおり、「移行処理前」(505)、「経路移行状態」(506)、「データ移行状態」(507)、「移行処理後」(508) の 4 状態がありそれぞれ図 13 (a)、図 13 (b)、図 13 (c)、図 13 (d) に対応する。

【0120】

「移行処理前」(505) では、移行先記憶サブシステム 103b は、移行元記憶サブシステム 103a に接続されている SAN 102a と接続されておらず、移行元記憶サブシステム 103a と移行先記憶サブシステム 103b とは物理的に切り離されている。また、ホスト 101 は、移行先記憶サブシステム 103b の SAN 102b と接続されていないので、ホスト 101 は移行先記憶サブシステム 103b にアクセスすることができない (図 13 (a))。

【0121】

「経路移行状態」(506) では、移行先記憶サブシステム 103b は、移行元記憶サブシステム 103a に接続されている SAN 102a と接続され、ホスト 101 から移行先記憶サブシステム 103b にアクセスでき、ホスト 101 の経路移行が可能な状態とする。また、移行元記憶サブシステム 103a は移行先記憶サブシステム 103b と SAN 102a を介して接続されているので、移行先記憶サブシステム 103b から移行元記憶サブシステム 103a に対して経路移行処理のためのアクセスをすることができる (図 1

3 (b) )。

#### 【0122】

ホスト101の経路移行が終了すると「データ移行状態」(507)となり、ホスト101から移行元記憶サブシステム103aへのアクセスを禁止しなければならない。このため、「データ移行状態」(507)では、ホスト101をSAN102bに接続し、ホスト101からSAN102を介して移行先記憶サブシステム103bへアクセスができるようにする。また、データを移行するためにSAN102aには移行元記憶サブシステム103aと移行先記憶サブシステム103bとが接続された状態を維持する(図13(c))。

#### 【0123】

データの移行終了後の「移行処理後」(508)では、移行先記憶サブシステム103bから移行元記憶サブシステム103aへのアクセスは不要となるため、移行元記憶サブシステム103a及びSAN102aを移行先記憶サブシステム103bから切り離す。ホスト101は、SAN102bに接続されているので、ホスト101からSAN102を介して移行先記憶サブシステム103bへアクセスすることができる(図13(d))。

#### 【0124】

なお、SAN102を使用せずに、直接ホスト101から移行先記憶サブシステム103bへ接続してもよい。この場合「経路移行状態」(506)ではSAN102aの代わりにホスト101と移行元記憶サブシステム103aの間、ホスト101と移行先記憶サブシステム103bの間、移行元記憶サブシステム103aと移行先記憶サブシステム103bの間の各々を接続するため3本の経路が必要である。「データ移行状態」(507)においては移行元103aと移行先記憶サブシステム103bの間、ホスト101と移行先記憶サブシステム103bの間の各々を接続する2本の経路が必要である。

#### 【0125】

SAN102を使用しない場合、移行先記憶サブシステム103bの接続性が低下するため、多くのホスト101や記憶サブシステム103が接続される大規模な情報処理システムを構成するときは、SAN102を使用するのが好ましい。

#### 【0126】

[ゾーニングによる例]

図14は、本発明の実施の形態のゾーニングによってアクセス制限をする情報処理システムの構成を表すブロック図である。

#### 【0127】

図14に示す実施の形態では、ネットワーク管理のために移行管理ホスト112にSAN管理API(Application Program Interface)1401を設ける。移行管理プログラム113は、データ移行処理の各状態に対するSAN102におけるアクセス経路の制限を、SAN管理API1401によって設定する。なお、他の構成は前述した実施の形態(図1)と同じなので詳細な説明は省略する。

#### 【0128】

図15は、本発明の実施の形態の情報処理システムの、ゾーニングによってアクセス制御を行う動作の説明図である。

#### 【0129】

ゾーニングは、ファイバチャネルにおけるポート固有の識別番号(WWN:World Wide Name)を用いて、複数のポートをグループ(ゾーン)に分ける。そして、SANを構成するファイバチャネルスイッチによって、アクセス元及びアクセス先を特定し、他のゾーンに対する又は他のゾーンからのアクセスを許可しないようにスイッチの切り替えを制御をするアクセス制限方法であり、ネットワーク(SAN)に備わったアクセス制限方法である。

#### 【0130】

移行処理は、図5に示すとおり、「移行処理前」(505)、「経路移行状態」(50

6)、「データ移行状態」(507)、「移行処理後」(508)の4状態がありそれぞれ図15(a)、図15(b)、図15(c)、図15(d)に対応する。

【0131】

ホスト101、移行元記憶サブシステム103a及び移行先記憶サブシステム103bがSAN102に接続されている。

【0132】

「移行処理前」(505)では、ホスト101と移行元記憶サブシステム103aとを含み、移行先記憶サブシステム103bを含まないゾーン1501が設定されているので、ホスト101からは移行先記憶サブシステム103bにはアクセスすることができない(図15(a))。

【0133】

「経路移行状態」(506)では、ゾーン1501に移行先記憶サブシステム103bを加えたゾーン1502を設定する。これによって、ホスト101からは移行元記憶サブシステム103a及び移行先記憶サブシステム103bの双方にアクセス可能となり、ホスト101の経路移行が可能な状態となる(図15(b))。

【0134】

ホスト101の経路移行が終了すると「データ移行状態」(507)となり、ホスト101から移行元記憶サブシステム103aへのアクセスを禁止しなければならない。このため、「データ移行状態」(507)では、ゾーン1502から移行元記憶サブシステム103aを除いたゾーン1503を設定する。また、移行元記憶サブシステム103aと移行先記憶サブシステム103bを含むゾーン1504を設定する。ゾーン1503はホスト101から移行先記憶サブシステム103bへのアクセスに使用され、ゾーン1504は移行元記憶サブシステム103aから移行先記憶サブシステム103bへのデータ移行に使用される(図15(c))。

【0135】

データの移行終了後の「移行処理後」(508)では、移行先記憶サブシステム103bから移行元記憶サブシステム103aへのアクセスは不要となるため、ゾーン1504を削除し、ホスト101から移行先記憶サブシステム103bへのアクセスのためゾーン1503のみが残り、ゾーン1503が継続して使用される(図15(d))。

【0136】

以上、ゾーニングによるアクセス制限の例を説明したが、同様にイーサネット(登録商標、以下同じ)におけるVLANによっても、アクセス可能なホストを制限することができる。

【0137】

[記憶サブシステムによって実現する例]

通常記憶サブシステム103は、LUセキュリティと呼ばれるボリューム105に対するアクセス制限の機能(以下、ストレージセキュリティ機能)を備えている。以下に説明する実施の形態では、記憶サブシステム103に備わるアクセス制限を使用してデータ移行処理のアクセス制御を実現する。

【0138】

図16は、本発明の実施の形態のストレージセキュリティ機能によってアクセス制限をする情報処理システムの構成を表すブロック図である。

【0139】

図16に示す実施の形態では、ネットワーク管理のために移行管理ホスト112に記憶サブシステム管理API(Application Program Interface)1601を設ける。移行管理プログラム113は、データ移行処理の各状態に対するSAN102におけるアクセスの制限を、記憶サブシステムAPI1601によって設定する。このアクセスの制限には、ホストに論理的に割り当てられた識別子(例えば、IPアドレス等のネットワーク上の識別子)や、ネットワークインタフェースに物理的に割り当てられた識別子(例えば、イーサネットにおけるMACアドレス、ファイバチャネルにおけるWWN)や、ネットワー

クインタフェースに論理的に割り当てられた識別子（例えば、iSCSI Name）を使用することができる。なお、他の構成は、前述した実施の形態（図1）と同じなので詳細な説明は省略する。

【0140】

図17は、本発明の実施の形態のストレージセキュリティ機能によってアクセス制御を行う動作の説明図である。

【0141】

移行処理は、図5に示すとおり、「移行処理前」（505）、「経路移行状態」（506）、「データ移行状態」（507）、「移行処理後」（508）の4状態がありそれぞれ図17（a）、図17（b）、図17（c）、図17（d）に対応する。

【0142】

ホスト101、移行元記憶サブシステム103a及び移行先記憶サブシステム103bがSAN102に接続されている。

【0143】

「移行処理前」（505）では、移行管理プログラム113は、移行元記憶サブシステム103aにはホスト101からのアクセスを許可し、移行先記憶サブシステム103bには外部からのアクセスを禁止するように設定する（図17（a））。

【0144】

「経路移行状態」（506）では、移行管理プログラム113は、移行元記憶サブシステム103aにはホスト101からのアクセスを許可し、移行先記憶サブシステム103bにもホスト101からのアクセスを許可するように設定する。また、経路移行中は、移行先記憶サブシステム103bから移行元記憶サブシステム103aへのアクセスが必要となるため、移行管理プログラム113は、移行元記憶サブシステム103aには移行先記憶サブシステム103bからのアクセスも許可するように設定する（図17（b））。

【0145】

ホスト101の経路移行が終了すると「データ移行状態」（507）となり、ホスト101から移行元記憶サブシステム103aへのアクセスを禁止しなければならない。このため、「データ移行状態」（507）では、移行管理プログラム113は、移行元記憶サブシステム103aにはホスト101からのアクセスを禁止し、移行先記憶サブシステム103bからのアクセスを許可するように設定する。移行先記憶サブシステム103bから移行元記憶サブシステム103aへのアクセスによってデータ移行が行われる。移行先記憶サブシステム103bには、「経路移行状態」（506）と同様に、ホスト101からのアクセスが許可される状態が設定される（図17（c））。

【0146】

データの移行終了後の「移行処理後」（508）では、移行先記憶サブシステム103bから移行元記憶サブシステム103aへのアクセスは不要となるため、移行管理プログラム113は、移行元記憶サブシステム103aに対する全てのアクセスを禁止する。また移行管理プログラム113は移行先記憶サブシステム103bに対してはホスト101からのアクセスを許可した状態を継続する（図17（d））。

【0147】

〔組み合わせ〕

図13～図17において、ネットワーク102、記憶サブシステム103におけるアクセス制限を実現する実施の形態を説明したが、ネットワーク102でのアクセス制限と、記憶サブシステム103でのアクセス制限とを組み合わせることによって、より堅牢なアクセス制限を実現することもできる。

【0148】

〔デバッグ状態付データ移行〕

経路移行処理においては、新経路の設定を確認する必要がある。そこで図3における経路移行処理（図3の303）においてデバッグモードを設けることがある。この場合は全ホスト101に対する経路変更を同時に実施する必要がある。

**【0149】**

図18は、本発明の実施の形態のデバッグモード付のデータ移行処理のフローチャートである。

**【0150】**

まず、移行前の状態において、前述した図3の初期設定(301)と同様に、移行管理プログラム113はデータ移行のための初期設定をする(1801)。

**【0151】**

初期設定の後、前述した図3の移行先記憶サブシステム追加処理(302)と同様に、移行先記憶サブシステム103bを情報処理システムに追加する(1802)。

**【0152】**

その後、データ移行に関係する全ホスト101に対して新経路を設定する(1803)。同時に、移行先記憶サブシステム103bを「デバッグモード」へと切り替える。

**【0153】**

「デバッグモード」では、移行元記憶サブシステム103aに対するアクセスを禁止する。移行先記憶サブシステム103bは、ホスト101からの書き込みデータを移行先記憶サブシステム103bへ格納する。書き込みデータに関してはビットマップに更新データの情報を記憶しておく。

**【0154】**

さらに、移行先記憶サブシステム103bは、ホスト101からの読み出し要求に関しては、ビットマップにおいて更新情報を検査し、移行先記憶サブシステム103bに書き込まれている場合には、移行先記憶サブシステム103bからデータを読み出す。一方、移行先記憶サブシステム103bに書き込みデータがない場合には、移行元記憶サブシステム103aからデータを読み出す。そして読み出したデータをホスト101へ送信する。

**【0155】**

その後、ホスト101にてテストプログラムを実行することによって経路確認を行って、新しく設定された経路が正しいか否かを判定する(1804)。

**【0156】**

経路が正しく設定されていなければ、ホスト101においてテストプログラムが実行されている間に、テストプログラムが移行元記憶サブシステム103aへアクセスを行おうとした場合、移行元記憶サブシステム103aはアクセス不可能となっているためエラーとなる(1805)。そして、エラー回数が規定回数内であるか否かを判定する(1806)。エラー回数に上限値を設けておき、規定回数内であれば再度設定を見直し(1807)、データを回復し(1808)、ステップ1803に戻って、テストプログラムを再度実行する。

**【0157】**

一方、エラー回数が規定回数を超えた場合、移行先記憶サブシステム103bは書き込みデータを破棄し、データを移行先装置追加時の状態(経路移行の前の状態)に戻すデータ回復処理を実行する(1809)。データ回復処理は、移行先記憶サブシステム103bへの書き込みデータを消去してもよいが、ビットマップに記録された更新情報を全て消去しても実現可能である。その後、この移行処理を処理を終了する。

**【0158】**

また、テストプログラムが正常に終了し、経路が正しく設定されていると判定されると、データを経路移行前の状態に戻すデータ回復処理を実行する(1810)。そして、データ移行処理を実行する(1811)。データ移行処理(1811)は、図3のデータ移行(304)処理と同様の処理を実行する。

**【0159】**

この移行処理の間(ステップ1803～1811)、移行先記憶サブシステム103bは移行元記憶サブシステム103aをアクセスすることができ、ホスト101は移行先記憶サブシステム103bをアクセスすることができる。

**【0160】**

図19は、デバックモード付のデータ移行処理（図18）の各状態における移行元記憶サブシステム103a、移行先記憶サブシステム103bの動作を示す。

**【0161】**

「デバック状態」（1906）では、ホスト101から移行元記憶サブシステム103aへはアクセス不可とし、移行先記憶サブシステム103bからの読み出しに関しては、未更新のデータは移行元記憶サブシステム103aから読み出し、更新済みのデータは移行先103bから読み出しホスト101へ送信する。ホスト101からの書き込み処理については、全てのデータを移行先記憶サブシステム103bへ格納する。

**【0162】**

なお、「移行処理前」「データ移行状態」及び「移行処理後」の各状態は、図5で前述したものと同じなので、その説明は省略する。

**【0163】**

以上説明したように、本発明の実施の形態では、データを移行する前に経路移行状態を設け、経路移行状態中でもホストコンピュータからのアクセスを可能としたので、データ移行時の可用性が向上する。

**【0164】**

具体的には、ボリューム単位でデータ移行処理をし、経路移行状態を設け、経路移行中のボリュームに係わるホストの経路を順次切り替える。そして、経路移行中のボリュームに複数のホストコンピュータがアクセスしている場合は、経路移行前のホストコンピュータは旧記憶サブシステムにアクセスし、経路移行後のホストコンピュータは新記憶サブシステムにアクセスする。このとき旧記憶サブシステムと新記憶サブシステムのデータ一貫性を保つために、旧記憶サブシステムに全ての更新情報が反映されるようにする。経路移行前のホストコンピュータは旧記憶サブシステムに書き込み／読み出し処理を実行し、読み出し処理時のデータは旧記憶サブシステムから読み出し、書き込み処理時のデータは旧記憶サブシステムに書き込む。また、経路移行後のホストコンピュータからの新記憶サブシステムに対するアクセスは、読み出し処理時は、新記憶サブシステムが旧記憶サブシステムから読み出したデータをホストへ送信する。書き込み処理時は新記憶サブシステムにデータが書き込まれる。

**【0165】**

このようにすることで、経路移行状態においても、経路未移行ホストから旧記憶サブシステムへのアクセス、及び、経路移行済ホストから新記憶サブシステムへのアクセスが同時に可能となる。また、経路移行状態において、旧記憶サブシステムに最新のデータが格納され、旧記憶サブシステムと新記憶サブシステムとのデータの一貫性が保証され、データ移行処理においてシステムの可用性を向上することができる。

**【0166】**

また、本発明に係る機能は、新記憶サブシステムのみによって実現可能であり、旧記憶サブシステムに新規の機能を追加する必要がない。

**【図面の簡単な説明】****【0167】**

【図1】 本発明の実施の形態の情報処理システムの構成を表すブロック図である。

【図2】 本発明の実施の形態の記憶サブシステムの構成図である。

【図3】 本発明の実施の形態の移行処理のフローチャートである。

【図4】 本発明の実施の形態の情報処理システムの動作の説明図である。

【図5】 本発明の実施の形態の移行処理の動作の説明図である。

【図6】 本発明の実施の形態の移行管理テーブルの説明図である。

【図7】 本発明の実施の形態の経路移行処理の動作を示すフローチャートである。

【図8】 本発明の実施の形態のデータ移行処理の動作を示すフローチャートである。

【図9】 本発明の実施の形態の経路移行状態とデータ移行状態における書き込み処理の概要の説明図である。



【図 10】本発明の実施の形態の書き込み処理の動作を示すフローチャートである。

【図 11】本発明の実施の形態の経路移行状態とデータ移行状態における読み出し処理の概要の説明図である。

【図 12】本発明の実施の形態の読み出し処理の動作を示すフローチャートである。

【図 13】本発明の実施の形態の情報処理システムの、物理的な接続によってアクセス制御を行う動作の説明図である。

【図 14】本発明の実施の形態のゾーニングによってアクセス制限をする情報処理システムの構成を表すブロック図である。

【図 15】本発明の実施の形態の情報処理システムの、ゾーニングによってアクセス制御を行う動作の説明図である。

【図 16】本発明の実施の形態のストレージセキュリティ機能によってアクセス制限をする情報処理システムの構成を表すブロック図である。

【図 17】本発明の実施の形態の情報処理システムの、ストレージセキュリティ機能によってアクセス制御を行う動作の説明図である。

【図 18】本発明の実施の形態の情報処理システムのデバッグモード付データ移行処理のフローチャートである。

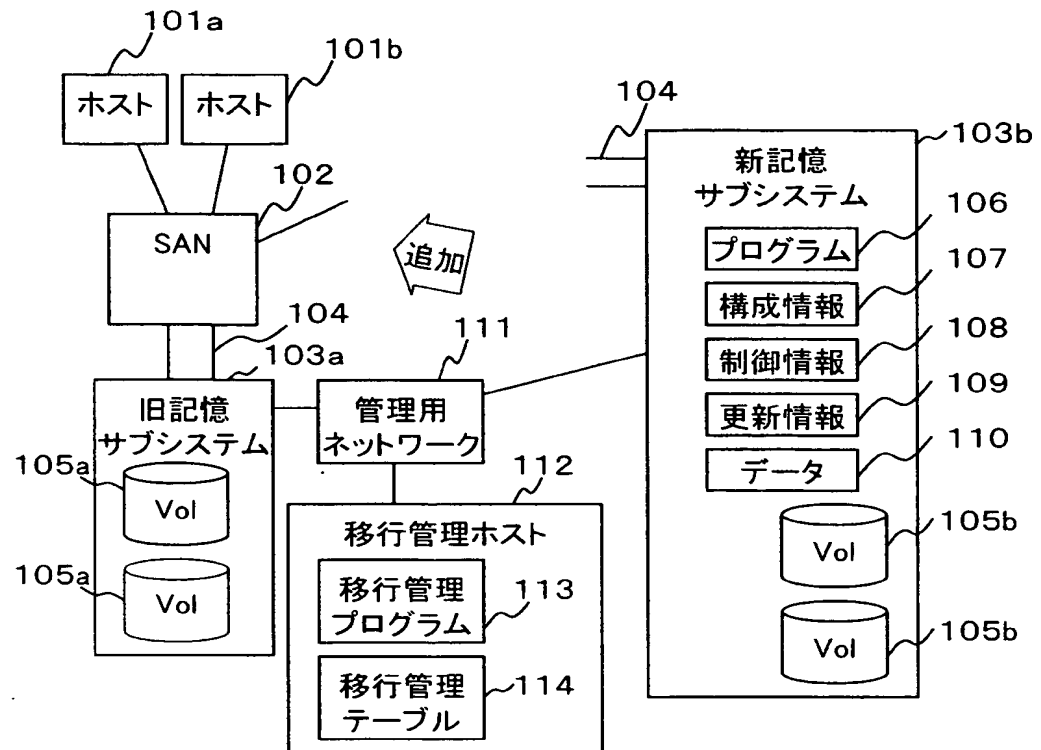
【図 19】本発明の実施の形態の情報処理システムのデバッグモード付データ移行処理の動作の説明図である。

【符号の説明】

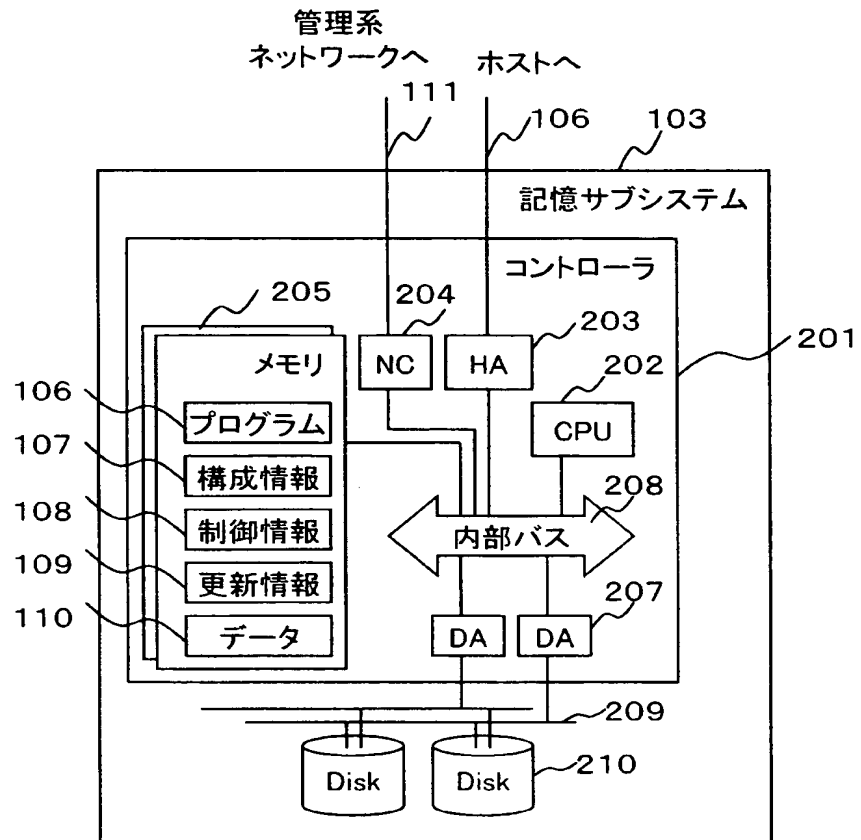
【0168】

- 101    ホスト
- 102    SAN
- 103    記憶サブシステム
- 104    インタフェース
- 105    論理ボリューム
- 106    プログラム
- 107    構成情報
- 108    制御情報
- 109    更新情報
- 110    データ
- 111    管理用ネットワーク
- 112    移行管理ホスト
- 113    移行管理プログラム
- 114    移行管理テーブル
- 201    コントローラ
- 202    CPU
- 203    ホストアダプタ
- 204    ネットワークコントローラ
- 205    メモリ
- 207    ディスクアダプタ
- 208    内部バス
- 209    ディスクインタフェース
- 210    ディスク装置

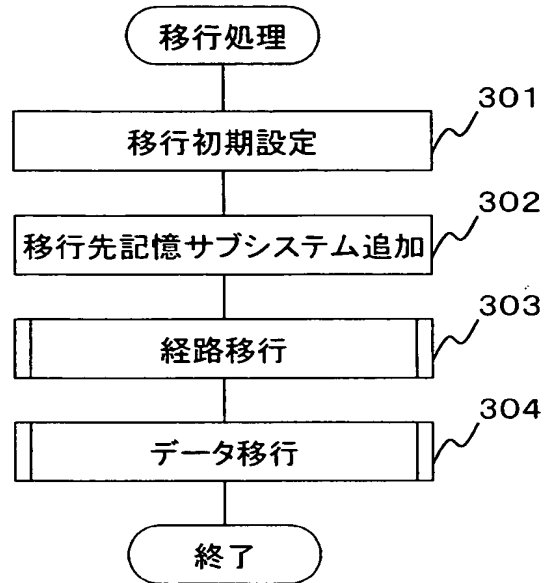
【書類名】 図面  
【図 1】



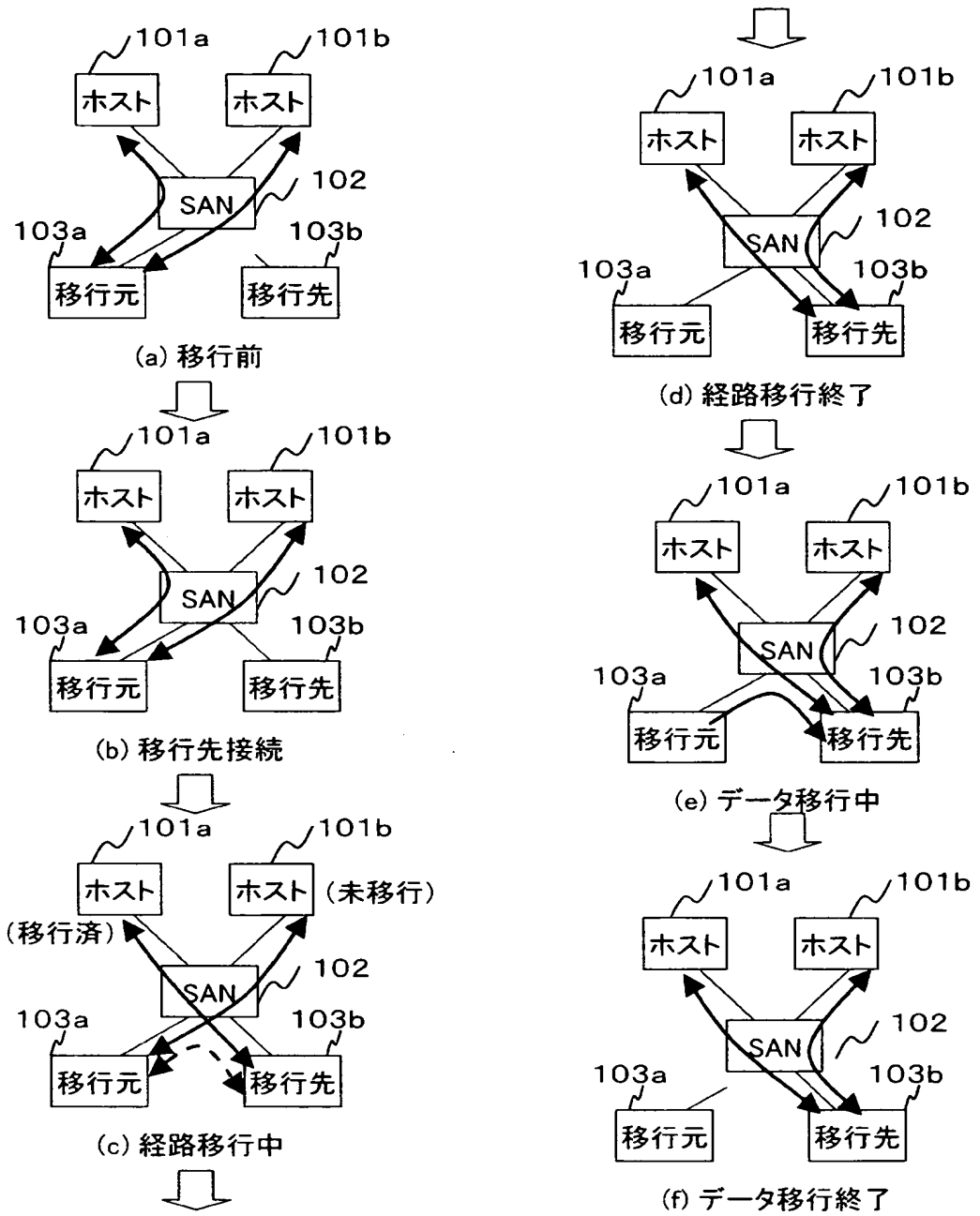
【図 2】



【図 3】



【図 4】



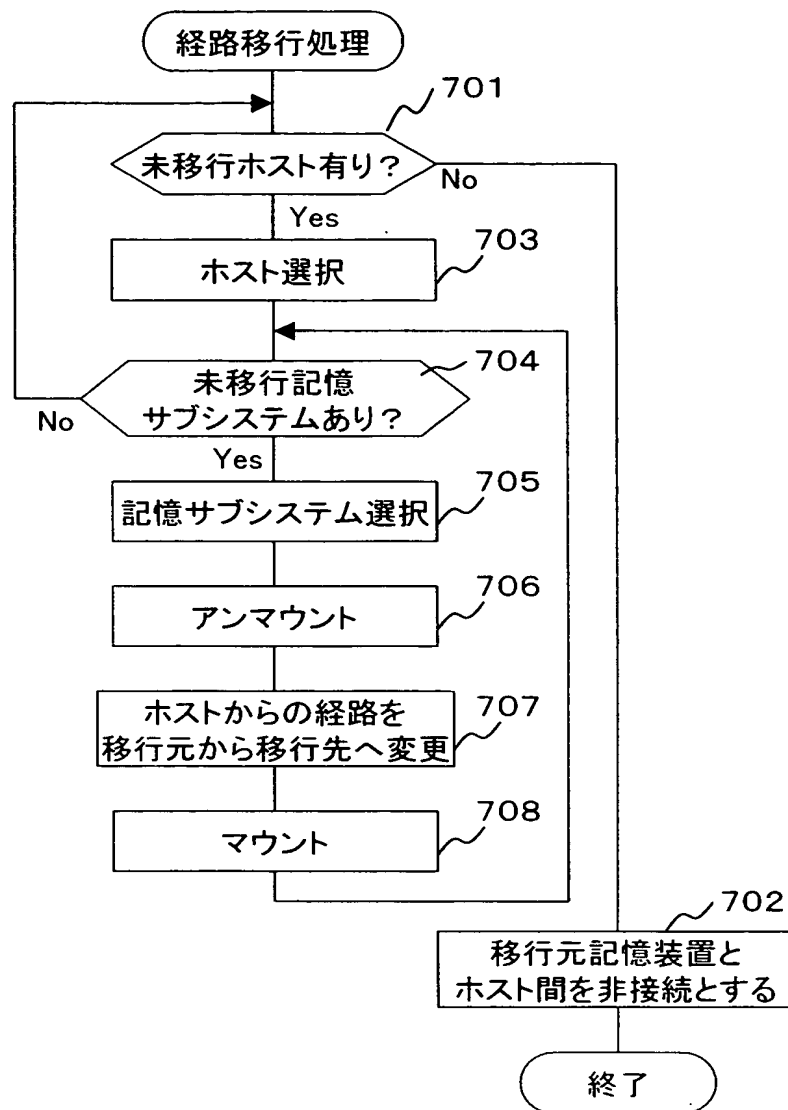
【図 5】

#	移行ステータス	移行元		移行先	
		読出	書込	読出	書込
1	移行処理前	元		アクセス不可	
2	経路移行状態	元		元→先	先→元
3	データ移行状態	アクセス不可		未領域: 元→先 済領域: 先	先
4	移行処理後	アクセス不可		先	先

【図 6】

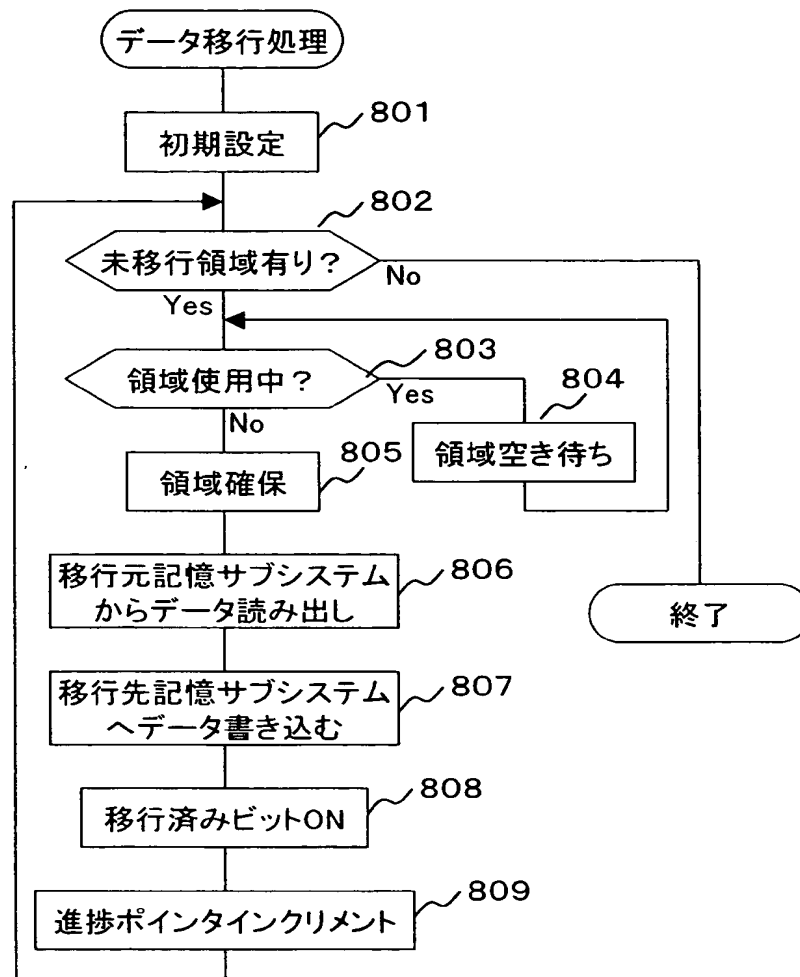
Gr. ID	Status	移行元 ID	移行先 ID	詳細
G00	移行済	S00-0000 S00-0001	S08-0000 S08-0001	/home /usr
G01	データ 移行中	S01-0000 S01-0001 S01-0002 S01-0003	S09-0000 S09-0001 S09-0002 S09-0003	/DBTable /log /backup0 /backup1
G02	経路 移行中	S02-0000 S02-0001 S02-0002 S02-0003	S09-0004 S09-0005 S09-0006 S09-0007	/dept0 /dept1 /dept2 /dept5
G03	未移行	S02-0004 S02-0005	S09-0008 S09-0009	/test0 /test1

【図 7】

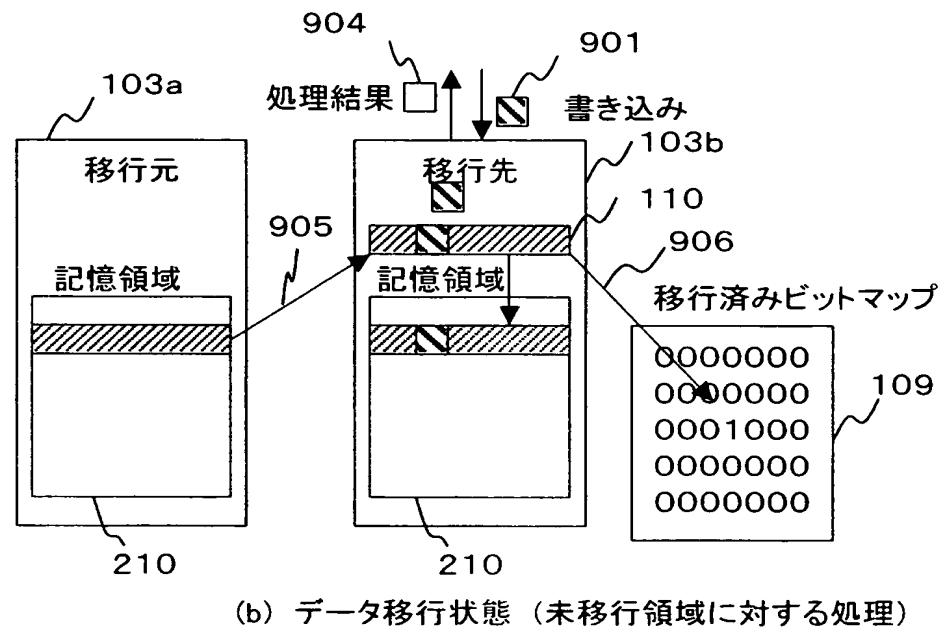
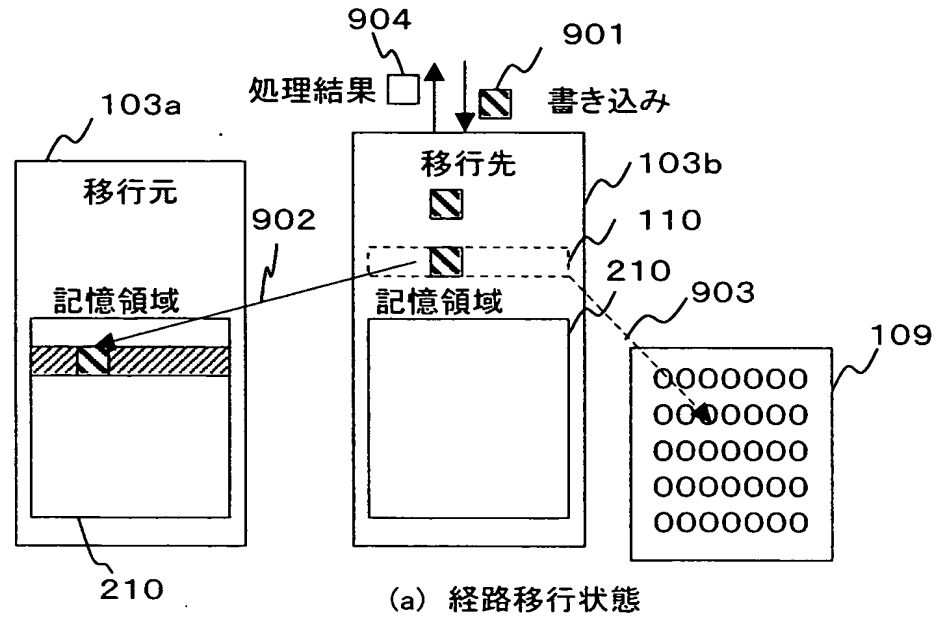




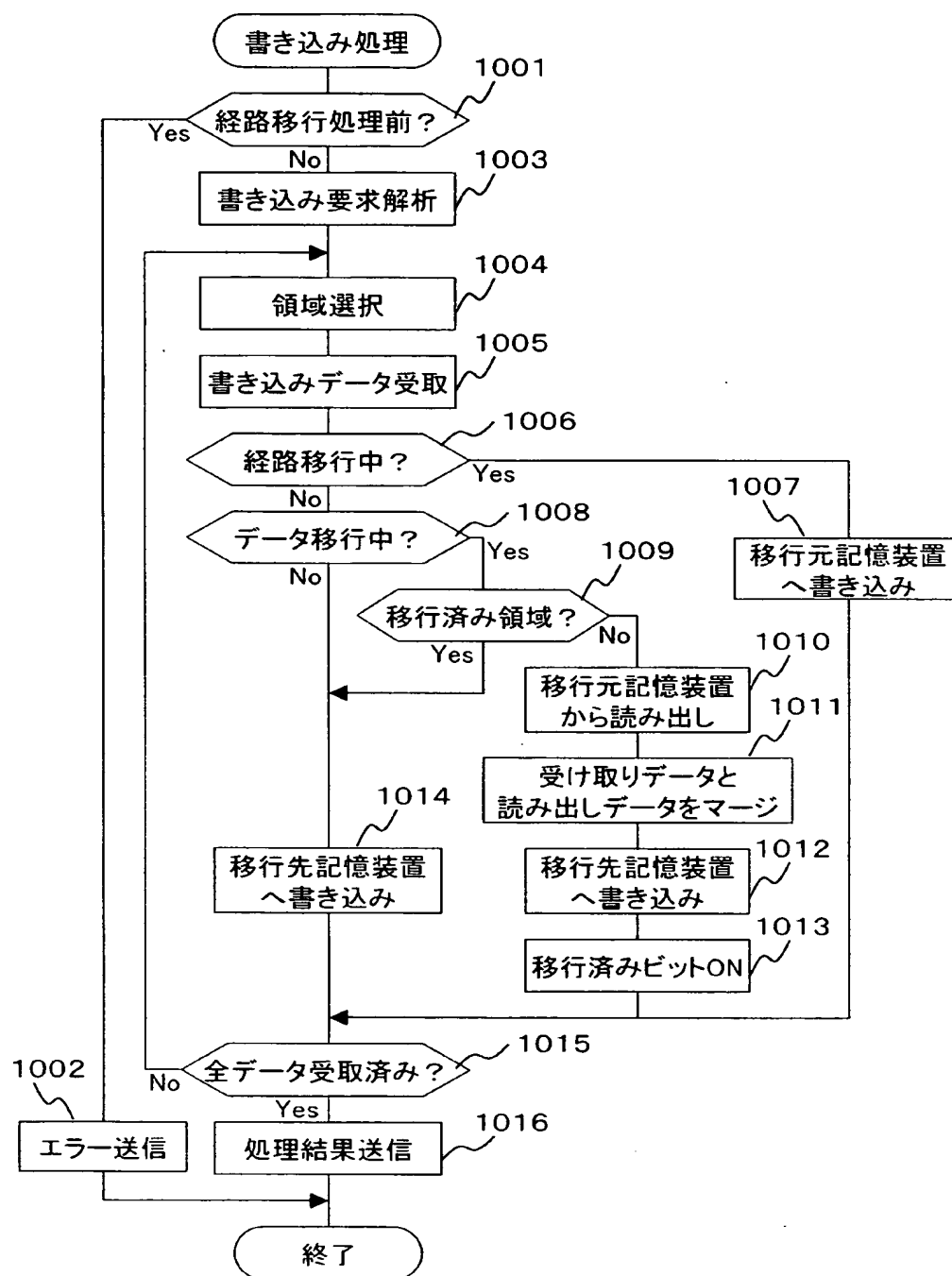
【図 8】



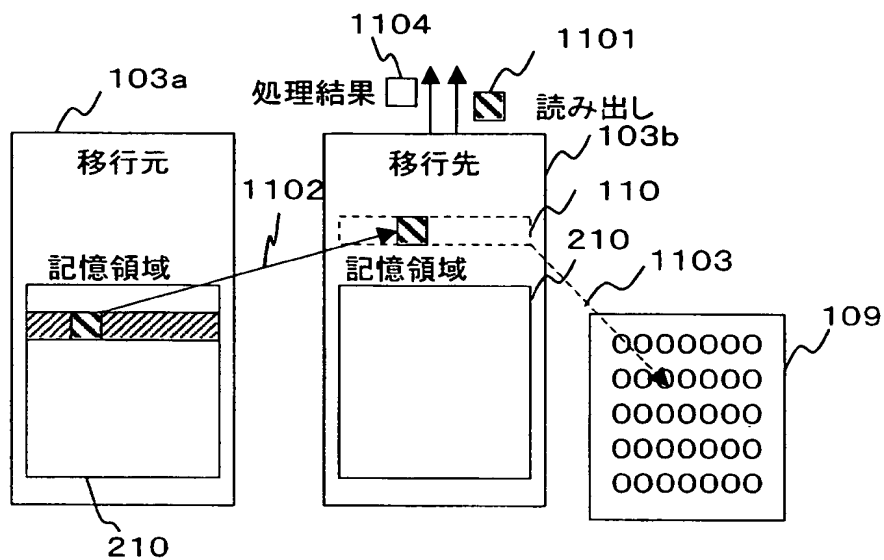
【図 9】



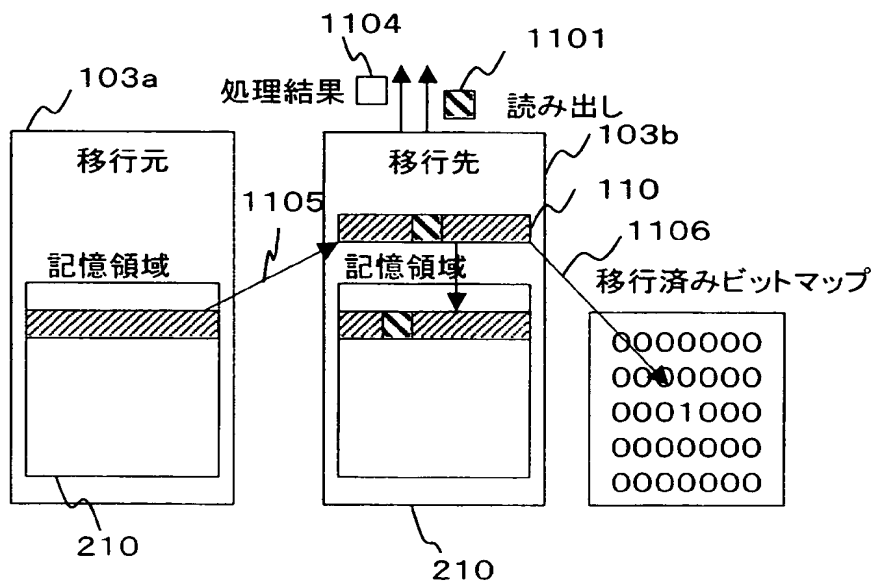
【図 10】



【図 11】

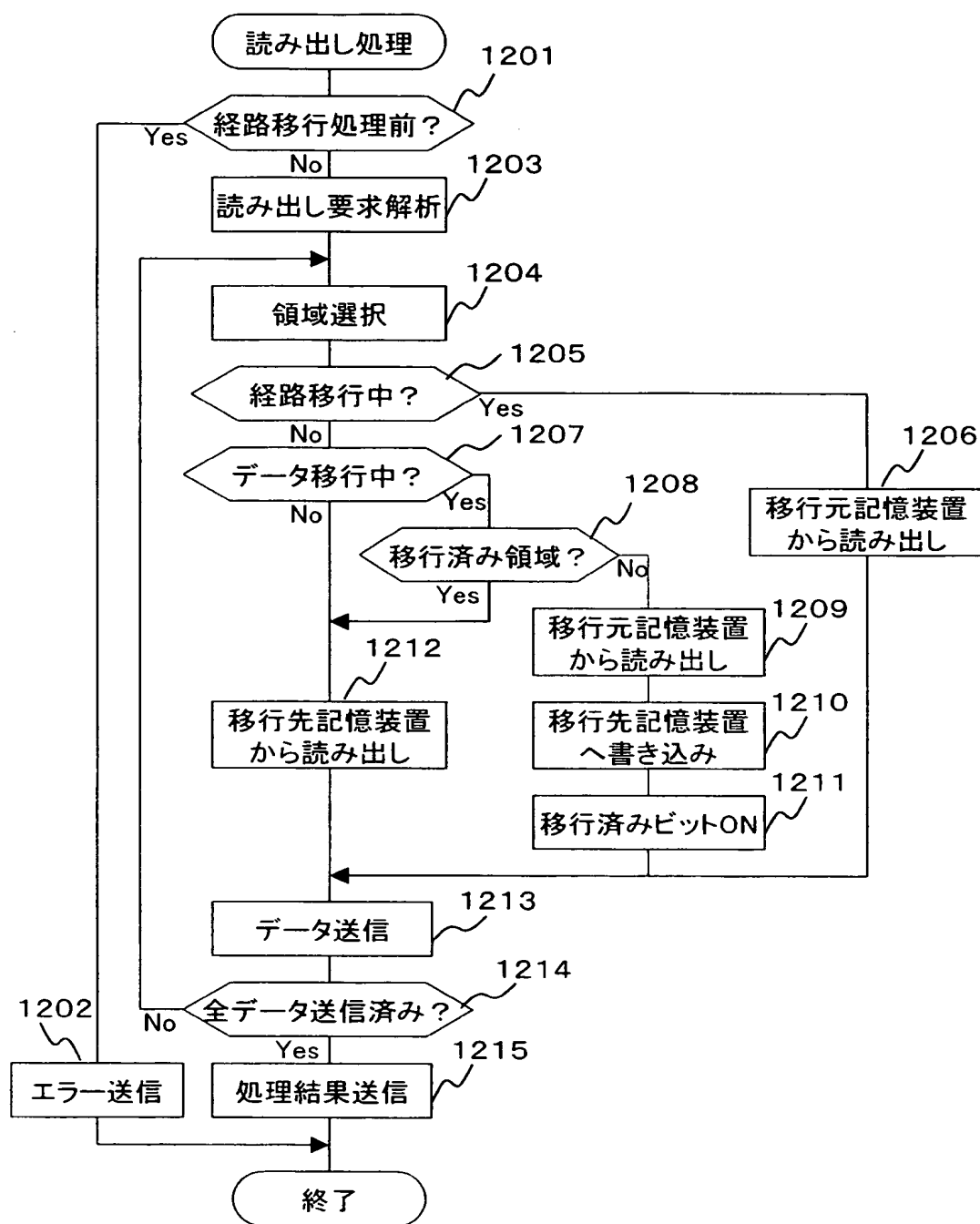


(a) 経路移行状態

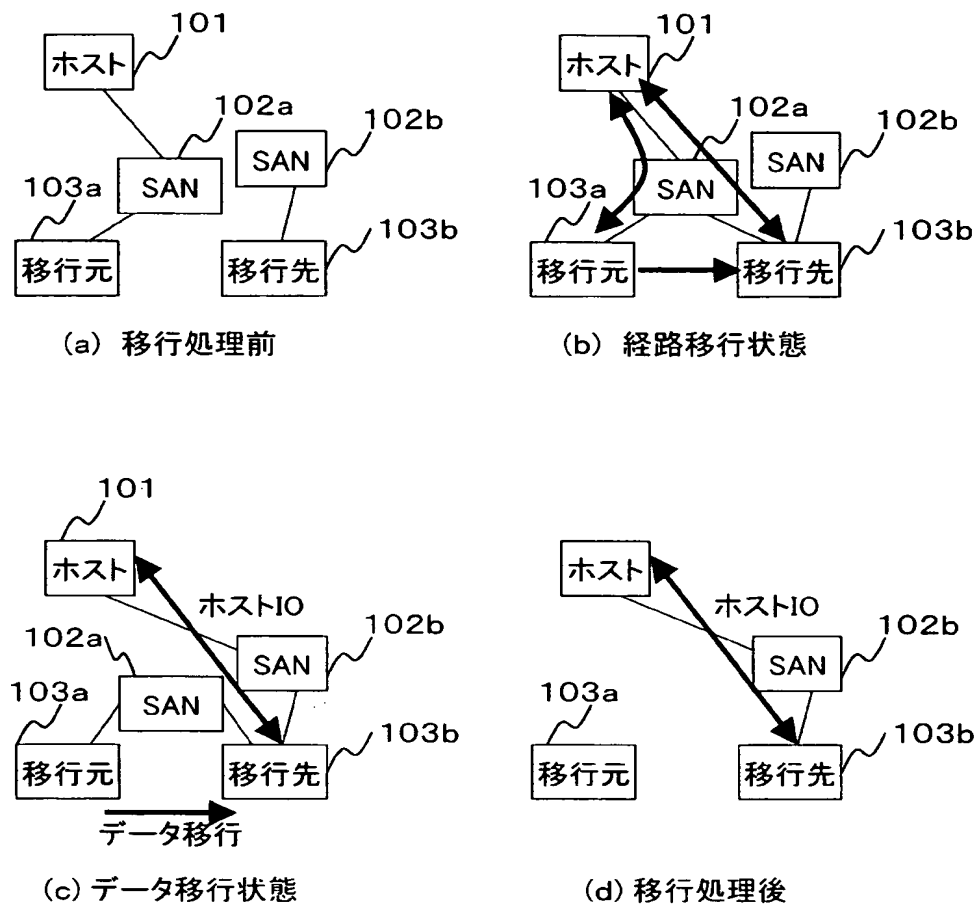


(b) データ移行状態 (未移行領域に対する処理)

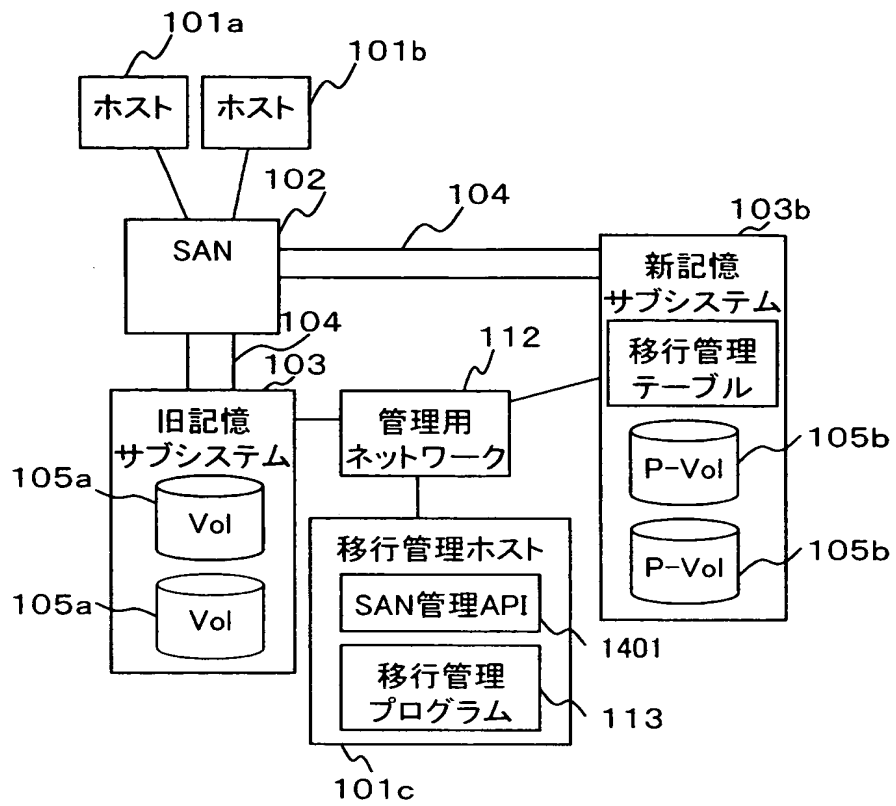
【図 12】



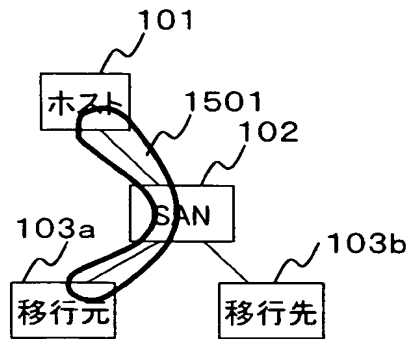
【図13】



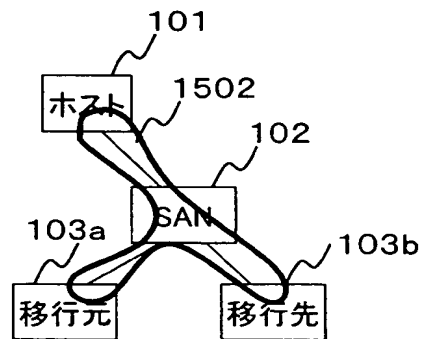
【図 14】



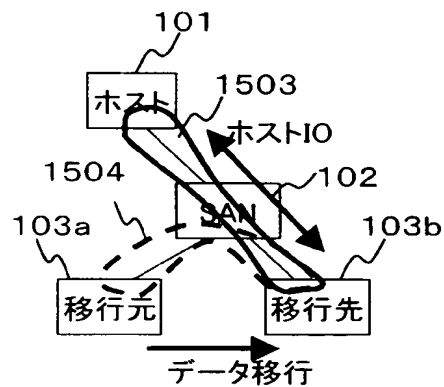
【図 15】



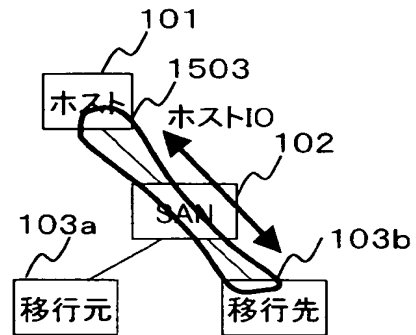
(a) 移行処理前



(b) 経路移行状態



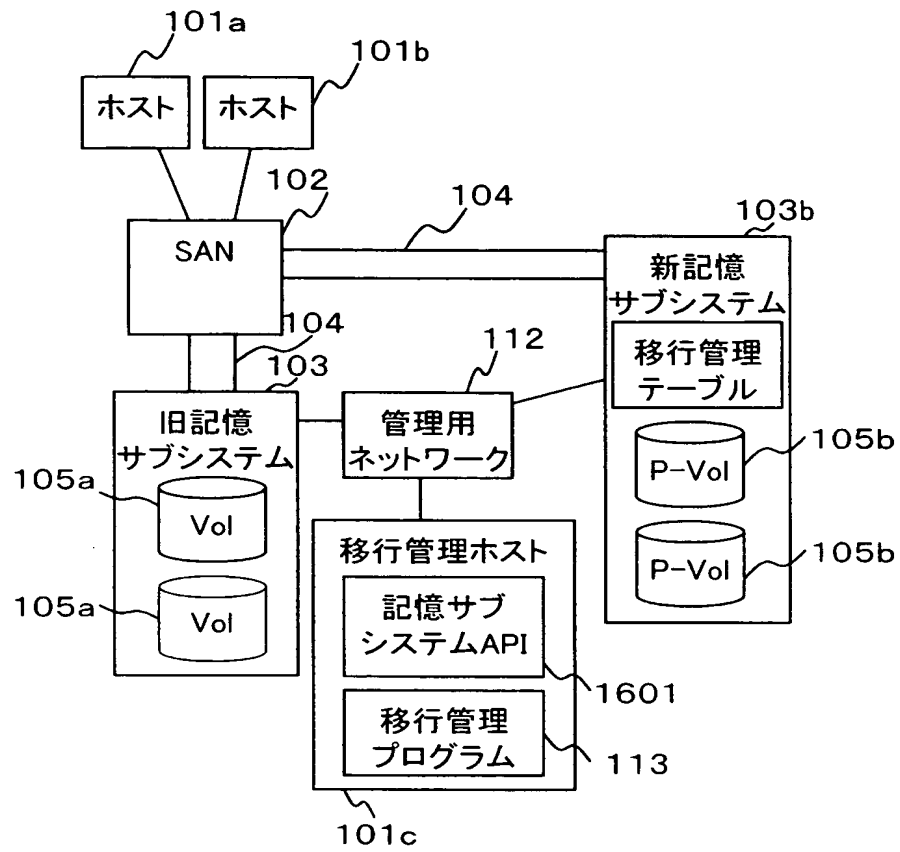
(c) データ移行状態



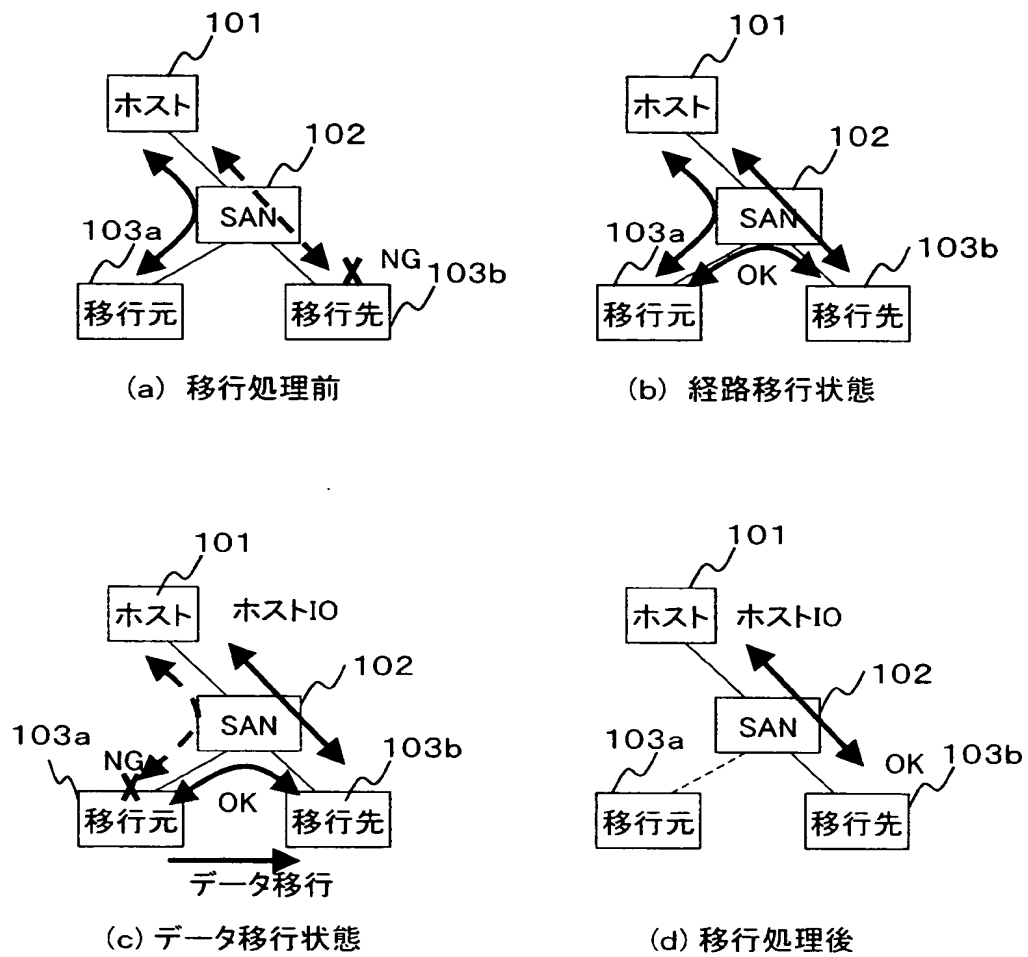
(d) 移行処理後



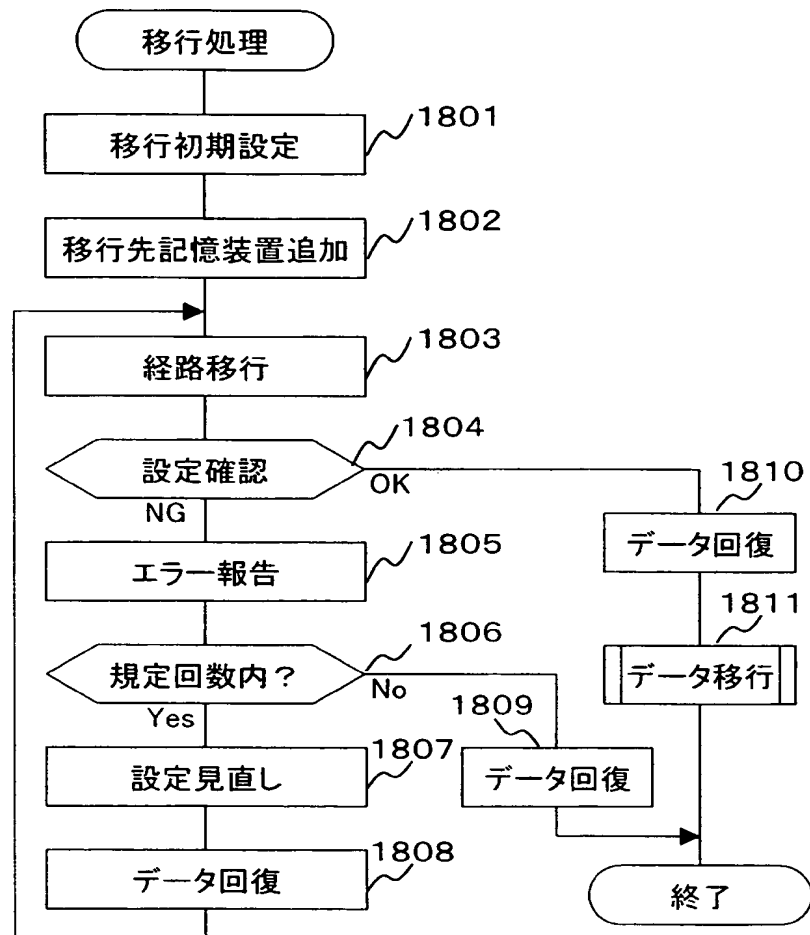
【図 16】



【図 17】



【図 18】



【図 19】

#	移行ステータス	移行元		移行先	
		読出	書込	読出	書込
1	移行処理前	元		アクセス不可	
2	デバッグモード	アクセス不可		未更新:元→先 更新:先	先
3	データ移行状態	アクセス不可		未領域:元→先 済領域:先	先
4	移行処理後	アクセス不可		先	先

**【書類名】 要約書****【要約】**

**【課題】**大規模、複雑なシステムにおいても可用性を低下させることのないデータ移行システムを提供する。

**【解決手段】**移行元記憶サブシステムに記憶されているデータを、移行先記憶サブシステムへ移行するデータ移行方法であって、前記移行元記憶サブシステムから前記移行先記憶サブシステムへデータを移行する前に、経路移行状態が設けられており、前記経路移行状態では、前記ホストコンピュータから前記移行元記憶サブシステム及び前記移行先記憶サブシステムへのアクセスを可能とするように経路を設定し、前記移行先記憶サブシステムは、前記ホストコンピュータからの書き込み要求に対し前記移行元記憶サブシステムにデータを書き込み、前記ホストコンピュータからの読み出し要求に対し前記移行元記憶サブシステムからデータを読み出して前記ホストコンピュータに送信する。

**【選択図】** 図 1

特願 2 0 0 4 - 0 0 5 6 3 5

出 願 人 履 歴 情 報

識別番号

[ 0 0 0 0 0 5 1 0 8 ]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所